SANDIA REPORT

SAND2005-5730
Unlimited Release
Printed November 2005
Reprinted January 2006

Risk Perception & Strategic Decision Making: General Insights, a New Framework, and Specific Application to Electricity Generation Using Nuclear Energy

Jeffrey D. Brewer

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2005-5730
Unlimited Release
Printed November 2005
Second Printing, January 2006

Risk Perception & Strategic Decision Making: General Insights, a New Framework, and Specific Application to Electricity Generation Using Nuclear Energy

Jeffrey D. Brewer
Risk & Reliability Department
Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-MS0748
Email: jdbrewe@sandia.gov

Abstract

The objective of this report is to promote increased understanding of decision making processes and hopefully to enable improved decision making regarding high-consequence, highly sophisticated technological systems. This report brings together insights regarding risk perception and decision making across domains ranging from nuclear power technology safety, cognitive psychology, economics, science education, public policy, and neural science (to name a few). It forms them into a unique, coherent, concise framework, and list of strategies to aid in decision making. It is suggested that all decision makers, whether ordinary citizens, academics, or political leaders, ought to cultivate their abilities to separate the wheat from the chaff in these types of decision making instances. The wheat includes proper data sources and helpful human decision making heuristics; these should be sought. The chaff includes ‘unhelpful biases’ that hinder proper interpretation of available data and lead people unwittingly toward inappropriate decision making ‘strategies’; obviously, these should be avoided. It is further proposed that successfully accomplishing the wheat vs. chaff separation is very difficult, yet tenable. This report hopes to expose and facilitate navigation away from decision-making traps which often ensnare the unwary. Furthermore, it is emphasized that one’s personal decision making biases can be examined, and tools can be provided allowing better means to generate, evaluate, and select among decision options. Many examples in this report are tailored to the energy domain (esp. nuclear power for electricity generation). The decision making framework and approach presented here are applicable to any high-consequence, highly sophisticated technological system.

Acknowledgements

I sincerely thank my colleagues at Sandia National Laboratories who generously allotted time to review and comment on various draft materials that eventually found their way into the present report. These individuals include: Mark S. Allen, David Borns, Nathan Brannon, J. Arlin Cooper, John Covan, Roger Cox, John Forester, Carmen Mendez, Randall Gauntt, Michael Hessheimer, Scott Jones, John Kelly, Ronald Knief, David Kunsman, Glenn Kuswa, Dwight Miller, David G. Robinson, Thomas Sanders, Ann Speed, Jenny Tobin, Marianne Walck, and Timothy Wheeler. Special thanks are also due to the formal reviewers of the near-final manuscript who provided insightful comments. Finally, I thank Mr. Mark Davenport at San Antonio College, San Antonio, TX, for introducing me to the work of Arnold B. Arons nearly fifteen years ago.

Contents

Figures

Tables

1. Introduction

It is suggested that all decision makers, whether ordinary citizens, academics, or political leaders, ought to cultivate their abilities to separate the wheat from the chaff in particular decision making instances. The wheat includes proper data sources and helpful human decision making heuristics; these should be sought. The chaff includes ‘unhelpful biases’ that hinder proper interpretation of available data and lead people unwittingly toward inappropriate decision making ‘strategies’; obviously, these should be avoided. It is further proposed that successfully accomplishing the wheat vs. chaff separation is very difficult, yet tenable. This report does not support an Orwellian[1]-type of decision making ‘wisdom’ that enables manipulation by an elite group with a specific agenda, but hopes to expose and facilitate navigation away from decision-making traps which often ensnare the unwary. Furthermore, it is emphasized that one’s personal decision making biases can be examined, and tools can be provided that allow better means to generate, evaluate, and select among decision options.

Understanding the factors that influence risk perception and strategic decision making is especially important when decisions regarding high-consequence, highly sophisticated technological systems are required. This report brings together insights regarding risk perception and decision making across domains ranging from nuclear power technology safety, cognitive psychology, economics, science education, public policy, and neural science (to name a few) and forms them into a unique, coherent, concise framework, and list of strategies to aid in decision making. Early drafts of this report focused exclusively on decision making[2] in the ‘energy domain’ as the Laboratory Directed Research and Development project supporting this work specifically dealt with energy and infrastructure systems analysis. Significant portions of this report are still tailored to that topic as risk perception and decision making tendencies ought to be presented in the specific milieu of a particular subject. The decision making framework and approach presented here are applicable to any high-consequence, highly sophisticated technological system. This report is intended to promote increased understanding of decision making processes and hopefully to enable improved decision making.

The order of exposition of this complex topic is as follows: first, a unique framework for understanding decision making is briefly presented along with a couple illustrative examples; second, critical thinking processes are defined and described which are considered foundational to proper decision making; third, the nature of a ‘technologically advanced’ society is briefly discussed;[3] fourth, a detailed discussion (including examples) of the foundations for the framework is presented; fifth, a step-by-

[1] In reference to George Orwell’s (1945) book *Animal Farm* in which an allegorical tale describes the success of a minority of power seekers in dominating the majority and imposing political and social tyranny (Orwell 1993).

[2] It should be noted that ‘risk perception’ is considered here as a subset of decision making as lingering risk perceptions involve choices made implicitly or explicitly regarding a particular risk-related object.

[3] Awareness of the basic characteristics of ‘technologically advanced’ societies (e.g., the U.S.) are immensely important for recognizing the increasing role of ‘trust’ by members of society in technologies which they do not, and cannot understand (i.e., due to the volume and complexity of such knowledge) versus times of old when people knew virtually everything about the tools and technologies on which they depended for survival.

step description of a recommended decision making approach is presented; sixth, a discussion of perceived risks and analyzed risks related to nuclear power is presented along with pointers to the proposed framework and decision making approach. Finally, a summary restates the essential components of the framework and approach along with an appeal for using this method for decisions involving high-consequence, highly sophisticated technological systems where time is available for thoughtful reflection.

2. A Unique Framework for Understanding Risk Perception and Decision Making

Every person makes many risk related decisions each day. What risks should be taken? What types of information should be used for determining which risks to take? How do people actually approach risky decisions? Maybe the first question to ask is: what is the good that people seek? Jeremy Bentham, an English philosopher who lived from 1748 to 1832 explained the concept of utility as "…that property in any object whereby it tends to produce benefit, advantage, pleasure, good or happiness…" (Berenstein 1998). Bernstein (1998) states:

> All of us think (of) ourselves as rational beings even in times of crisis, applying the laws of probability in cool and calculated fashion to the choices that confront us. We like to believe we are above-average in skills, intelligence, farsightedness, experience, refinement, and leadership. Who admits to being an incompetent driver, a feckless debater, a stupid investor, or a person with an inferior taste in clothes? Yet how realistic are such images? Not everyone can be above average. Furthermore, the most important decisions we make usually occur under complex, confusing, indistinct, or frightening conditions. (p. 269)

When approached with an opportunity to take a course of action there is a continuum of consequences along which the decision scenario may be placed. For decisions involving actions having low to moderate consequences (e.g., which route to take to work, how much to eat during lunch, etc.), it does not matter too much which decision is made or the ensuing actions taken in one of these particular circumstances. Also, there are frequently encountered situations in which the 'right' decision to be made rapidly emerges (e.g., accidentally cutting your finger severely while chopping carrots, which immediately prompts you to cover the wound, apply pressure, and then seek medical attention). This process of rapidly recognizing the right decision/action pairs has led to the recognition-primed decision (RPD) model of decision making (Klein 1998)[4]. The RPD model is a natural extension of the phenomenon of *satisficing*, which involves selecting the first option that works (Simon 1957).

Satisficing is often much more expedient than optimization, since optimization requires identification of the 'best' option, which requires *explicit* identification of the criteria needed to evaluate the options[5]. For example, if a person is trying to fix a home

[4] See Klein (1998) for a good review of this model and how it can account for various decision making processes.

[5] Explicit identification of decision criteria is non-trivial; experts who perform very well in complex situations often have tremendous difficulty articulating in a detailed fashion what led them to select a 'successful' option. Refer again to Klein (1998).

appliance in their workshop, they will likely continue to analyze the problem and tinker until they find a solution that works (i.e., enables successful operation of the appliance). Continuing to search for a better option would not be likely to occur unless the first option in fact proved to be ineffective at achieving operational status.

The satisficing approach turns out to be quite helpful for many decision/ action situations. But is it ideal for all situations? Certainly there are many decision making domains involving high consequence, highly complex phenomena which ought not to be addressed without careful reflection and deliberation. Figure 2.1 shows a range of systems along a continuum of the level of care required, consequences of failure, and tolerance of failure. The decision regarding the type of energy source (or mix of sources) that should be used by the United States for supplying base load electricity, ensuring sustainability for growing electricity needs and ensuring safety of both the people and the environment is a critical, complex (i.e., multiple-attributes that are often in competition) decision making arena located near the right side of Figure 2.1. Making the correct decision for such important matters requires prudent decision making skills. This section briefly introduces a range of decision making biases and then presents a unique framework which can serve as an aid to decision making.

Figure 2.1. Qualitative assignment of several decision domains along three relevant dimensions. Adapted from Covan and Cooper (2004).

When approaching any decision, especially one that involves complex, high-consequence systems, an important consideration is to consciously review what is known and suspected about the domain of interest and to probe areas where unknown factors may lurk. Several months after the terrorist attacks on the U.S. on September 11, 2001, Secretary of Defense Donald H. Rumsfeld provided a description of imperfect knowledge while responding to a question about what the U.S. knew about weapons of mass destruction and terrorist support activities in the country of Iraq (Rumsfeld 2002):

The Unknown
As we know,
There are known knowns.
There are things we know we know.
We also know
There are known unknowns.
That is to say
We know there are some things
We do not know.
But there are also unknown unknowns,
The ones we don't know
We don't know.

The above quotation, criticized by some as a bit wordy and repetitive, is quite correct. That is, the accumulation of knowledge about any complex topic is a challenging endeavor that involves an indirect journey out of ignorance. Along that journey one makes observations and proposes inferences in the face of many uncertainties. Some of these uncertainties can eventually be categorized into 'known unknown'–areas of uncertainty, which are known to exist, but there is no clear way or no practical way to increase knowledge in those areas. Always lurking are the 'unknown unknowns'–gaps in knowledge that are critical to understanding the complex topic, but have not been identified at all. Figure 2.2 provides possible visualizations of one's state of knowledge regarding a system with conceptual boundaries established by defined system goals[6].

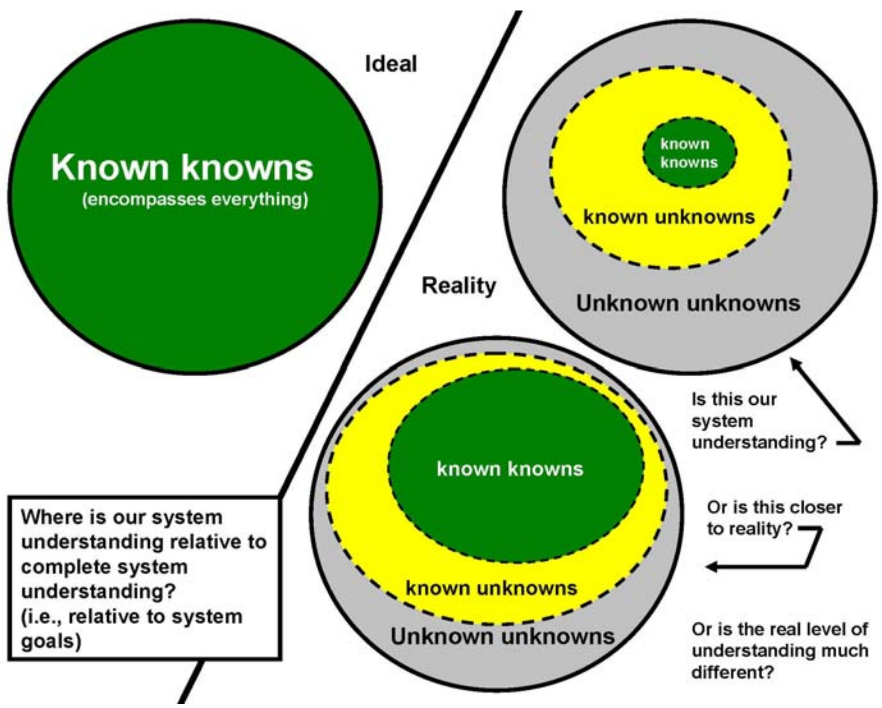

Figure 2.2. One's state of knowledge regarding a system.

[6] The reader should note that very often the system goals or requirements initially chosen are not nearly ideal. Iterating through several versions of system goals over a significant period of time should always be anticipated for high-consequence, technologically advanced systems.

In the community of risk and reliability professionals, two words are increasingly used to dichotomize uncertainties as being either *aleatory* or *epistemic* in nature (Hora 1996; Parry 1996; Zio and Apostolakis 1996). Aleatory uncertainty involves events or phenomena that occur in a random or stochastic fashion. A simple example of aleatory uncertainty would be random variations introduced when a rubber ball is dropped twice from the same location. Newtonian laws of motion and laws of energy conservation are very effective for enabling predictions of general trajectories the ball may follow and heights of intermediate bounces before the rubber ball comes to rest on the floor. But subtle, random variations will occur in how the ball is released, the specific portion of the ball that hits the floor first, imperfections in the spherical shape of the ball, etc. These are aleatory uncertainties in the rubber ball drop system. Another example of aleatory uncertainty would be anticipating the type of weather conditions present in the vicinity of a chemical plant at the time a pressure vessel explodes and releases chlorine gas.

Epistemic uncertainty involves one's confidence in the correctness of their observations and inferences concerning a system, and confidence in derived models or representations of that system. Both known unknowns and unknown unknowns fall into the epistemic category. A 'known unknown' example of epistemic uncertainty, using the rubber ball drop system could be the general shape of trajectories taken by the ball on successive bounces when all that is needed is the average time taken for the ball to come to rest on the floor. Detailed analysis could be performed to turn this known unknown into a known item if desired. Epistemic 'known unknown' uncertainties include modeling structure and specific model parameters. An example of an 'unknown unknown,' never conceived by the person analyzing the system, might be the occurrence of a rubber ball splitting into two pieces upon contact with the floor due to defects caused during the rubber ball manufacturing process. Epistemic 'unknown unknown' uncertainties include modeling structure and modeling completeness mistakes and omissions.

Understanding the general types of uncertainties involving complex, high-consequence domains is helpful. But, how exactly does someone move along the path from a state of ignorance or impoverished knowledge to a state of adequate or extensive knowledge? In this case adequate or extensive refers to the amount and quality of knowledge required to make an appropriate decision regarding high-consequence, advanced technology. To answer this important question, we begin with summarizing a unique framework.

This report proposes a unique framework for understanding key aspects of risk perception and decision making. The proposed taxonomy of biases begins with identifying the categories of **normative knowledge**, **availability**, and **individual specific** attributes into which the many biases/tendencies are grouped (see Figure 2.3). **Normative knowledge** involves a person's skills in combinatorics[7], probability theory, and statistics. Research has shown that training and experience in these quantitative fields can improve one's ability to accurately determine event likelihoods. Those trained in statistics tend to seek appropriate data sources when assessing the frequency and severity of an event. The **availability** category of biases includes those which result from the structure of human cognitive machinery. Two examples of biases in the availability

[7] Combinatorics is the branch of mathematics concerned with counting, arranging, and ordering (Larsen and Marx 2001).

category include the anchoring bias and the retrievability bias. The anchoring bias causes a decision maker to bias subsequent values or items toward the first value or item presented to them. The retrievability bias refers to the bias that drives people to believe those values or items which are easier to retrieve from memory are more likely to occur. **Individual specific** biases include a particular person's values, personality, interests, group identity, and substantive knowledge (i.e., specific domain knowledge related to the decision to be made). **Critical thinking skills** are also offered as foundational for competent risk perception and decision-making as they can mute the impact of undesirable biases, regulate the application of one's knowledge to a decision, and guide information gathering activities. These critical thinking skills will be defined and discussed in detail in section three of this report. In addition to borrowing insights from the literature domains mentioned in the introduction, the formal decision-making approach supported in this report incorporates some methods used in multi-attribute utility theory.

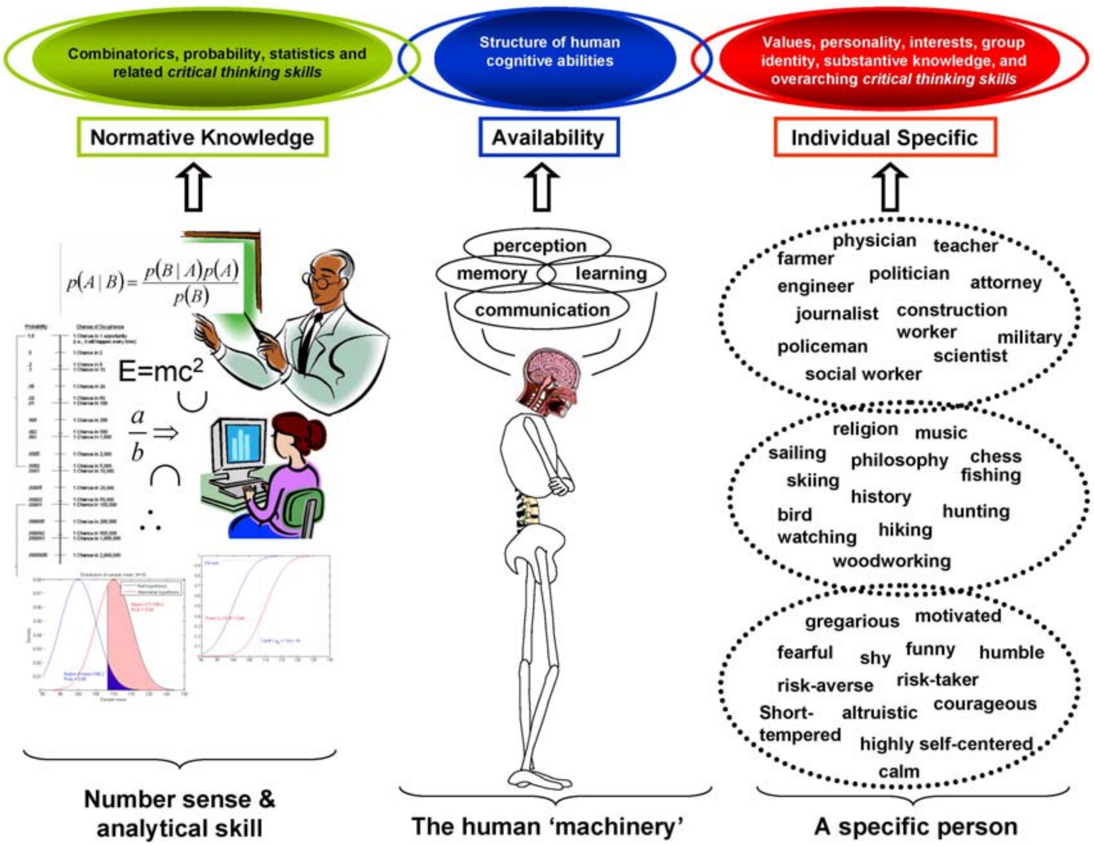

Figure 2.3. Main headings/categories of the taxonomy used by the author to group decision making biases or tendencies. The *critical thinking skills* referred to above (i.e., above the headings of **normative knowledge** and **individual specific** attributes) are defined and discussed at length in section three of this report.

A simple listing of the many biases discussed in detail in this report is provided in Figure 2.4. The biases are placed within the categories of the unique framework to which they are closely related. For example, the first bias listed in the normative knowledge category is *insensitivity to sample size.* Failing to temper the decision weight of statistical parameters based on the size of the sample from which the statistics are obtained indicates that people tend to not understand and apply insights from probability theory and statistics appropriately. One who is experienced in the normative domain will be much better equipped to identify and avoid such a pitfall. A second example of the bias grouping includes the first bias listed in the availability category (i.e., *anchoring effect*). People exhibit a strong tendency to overweight the first item they are shown or interact with in decision making settings. Improving knowledge regarding this effect and, ideally, increasing knowledge of human cognitive processes will improve identification and avoidance of this bias. The final example involves the first bias listed in the individual specific category (i.e., *loss aversion*). Loss aversion tendencies in decision making can vary tremendously from one person to another. Improving explicit self-knowledge regarding one's values, personality, interests, group affiliations, and areas of expertise can help explain why one is willing to accept particular levels of loss (i.e., how much risk they are willing to take) in a decision making setting.

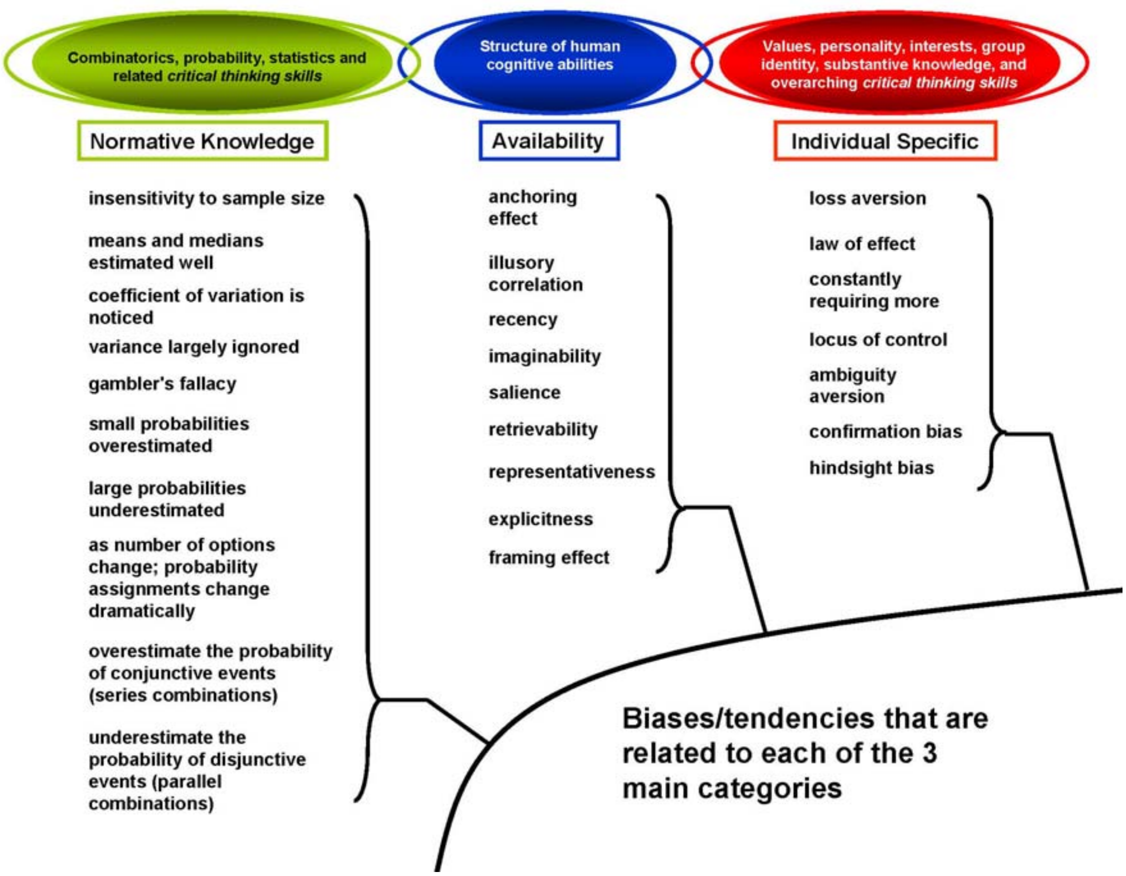

Figure 2.4. Listing of the decision making biases/tendencies related to each of the main categories in the unique framework. These lists of biases should be understood as persistent, undesirable decision making tendencies. It is proposed that the biases may be identified and mitigated by a decision maker by strengthening normative knowledge, understanding aspects of human cognition, increasing self-knowledge, and then following the approach described in section six of this report.

A comprehensive presentation of the unique framework proposed in this report for understanding key aspects of risk perception and decision making is provided in Figure 2.5. This figure shows not only the main categories and related biases, but also contains other novel features *introduced by the author*. The unique aspects of this framework include the following: (1) the definition and selection of three primary categories into which previously investigated biases are ordered; (2) hypothesized dependencies among biases both within and between categories; and (3) degrees to which these biases may be mitigated via focused effort by a particular decision maker. The evidence presented later in this report regarding many biases is intended to support the face validity of the unique framework and the associated decision-making improvement strategies.

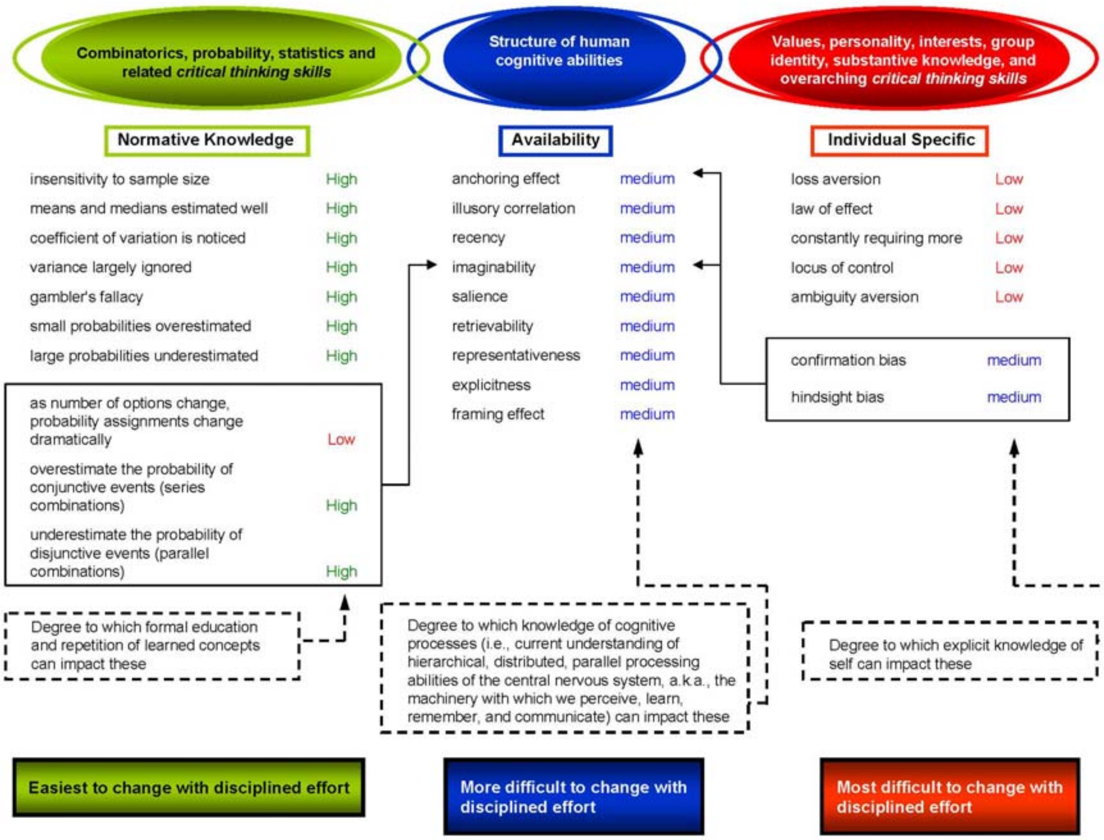

Figure 2.5. Biases/tendencies that researchers have identified, the ways in which these may be related and the ease with which one may become a better decision maker. The unique aspects *introduced by the author*, following a review of relevant data, include 1) the three primary categories into which previously investigated biases are ordered, 2) strong interdependencies hypothesized between biases as indicated by solid lines with arrows, 3) degrees to which these biases may be mitigated via focused efforts by a specific decision maker indicated by dashed lines with arrows; high, medium, and low ratings; and the green, blue, and red boxed items. This report is intended to present strong arguments supporting the face validity of the unique framework above and the associated decision-making improvement strategies.

A powerful example illustrating a number of biases in the framework includes the following hypothetical scenario: a respected energy expert who specialized in electricity production using fossil fuels was confronted with risk data showing the outstanding safety record and cost competitiveness of nuclear energy relative to electricity from fossil fuels (i.e., dramatically higher safety and comparable or even lower costs–without the emission of airborne pollutants or greenhouse gases). Instead of acknowledging the data and allowing himself to be open to the possibility that nuclear power might be as viable an energy source as fossil fuel, he recalled watching the Three Mile Island Unit 2 (TMI-2) accident unfold on television in 1979 and expressed how dangerous the situation was and how close to disaster we had come. He could recall salient images of people racing to leave the nearby communities and the recommendation that all pregnant women be evacuated within a five mile radius. He could also recall contradictory information being presented by 'experts' regarding the severity of the event. Following this comment about TMI-2 he was presented with data showing how there was essentially no danger to the public from that accident and also with data showing how slight the dangers would likely have been from even a severe 'reactor core meltdown.' In response, the fossil fuel expert stated, "yes, but nuclear energy was supposed to be the energy panacea–it was supposed to be too cheap even to meter–free electricity for all[8]."

Another example of biases is shown in this hypothetical scenario: a civil engineer expressed a conviction that thousands of people died of acute radiation exposure following the Chernobyl Unit 4 explosion in 1986 and that hundreds of thousands more people will eventually die of cancers caused by lower levels of radiation exposure from the accident. She was confronted with data from numerous investigations conducted by a host of nations revealing only 31 fatalities due to acute radiation exposure (mostly fire fighters) and 10 fatalities due to thyroid cancer years later (out of 800 children treated for thyroid cancer[9]). Data were also shown describing the levels of radiation exposure to various populations in the former Soviet Union and other countries due to the accident. This was coupled with data describing tremendous variations in radiation exposure due to natural sources around the world and the lack of epidemiological connections to cancer rates. She recalled hearing about thousands of people storming hospitals in the city of Kiev to be tested for radiation sickness; recalled a speaker at an anti-nuclear power rally[10] in 1972 declaring over a loudspeaker that, "…there's no such thing as a safe dose of radiation"; and distantly recalled the stories of mass graves during Stalin's rule.

Following these reflections she was shown data that explained that the frightened people going to the hospitals were found to have been unexposed to any significant amounts of radiation relative to background levels. She was also shown videos of the massive evacuation of all residents from Pripyat (the small city nearest Chernobyl; population approximately 57,500) and the extensive decontamination activities that were performed with technical cooperation by other countries. Data were also presented that showed

[8] The origin of producing electric energy that is too cheap to meter goes back at least to John von Neumann in 1955 as noted in Tenner (1996).

[9] The incidence of thyroid cancer in children in the affected areas was essentially zero before the accident, thus all the cases were attributed to radioactive iodine from the Chernobyl release.

[10] She reminisced that the atmosphere of the rally was exciting and the music was great–those were good times.

some health benefits to low exposures of radiation. The civil engineer responded, "yes, but didn't the Soviets use chemical agents on their own military during training exercises? Their lack of concern for people must have led to mass casualties, which they probably buried with the trees they cut down. As for the hundreds of thousands yet to die–there's no such thing as a safe dose of radiation."

The preceding examples reveal many decision making biases (e.g., confirmation bias, availability biases, possible overestimation of small probabilities, overestimation of the probability of conjunctive events, possible identification with anti-nuclear groups, etc.) that propel the decision maker toward a non-rational decision, that is, a decision in which extensive, valid data is ignored. Unfortunately, people's decision making biases tend to be ubiquitous, even for extremely well-educated individuals[11]. It is the position of this report that many people can learn how to mitigate such biases for a particular category of decisions. Techniques to counter the undesirable tendencies include a strong commitment to reflect upon one's potential biases in a specific decision making situation, to make decisions using the most valid quantitative data available, and to carefully map out what one considers important in the decision-making setting. Due to the importance of making appropriate decisions regarding high-consequence, technologically advanced, multi-attribute systems (e.g., electricity generation infrastructure for the U.S. population), a systematic approach is suggested–and one is presented in this report.

3. Critical Thinking Processes

> The simple but difficult arts of paying attention, copying accurately, following an argument, detecting an ambiguity or a false inference, testing guesses by summoning up contrary instances, organizing one's time and one's thought for study–all these arts…cannot be taught in the air but only through the difficulties of a defined subject; they cannot be taught in one course in one year, but must be acquired gradually in dozens of connections.
>
> Attributed to Jacques Barzun, cited in Arons (1990)

This section presents ten critical thinking processes or skills that are deemed essential to enable one to properly understand their interactions with their environment and with other people; therefore, they are foundational for decision making activities. The list is not intended to be complete or exhaustive, yet it is intended to define key concepts which are often thrown out casually (i.e., rarely defined) by educators and other commentators lamenting the decline of 'thinking-reasoning capacities,' 'critical thinking skills,' 'higher order thinking skills,' or 'critical thought.' The material for this section borrows very heavily from *A Guide to Introductory Physics Teaching*, published in 1990 by the late Arnold B. Arons.

Arons spent decades studying the teaching and learning of science and math by grade school children, and both science and non-science majors in college; he even included university professors in his research. Much of the research of Arons and his colleagues

[11] In this case *well-educated* implies one who has received extensive formal education but may not have extensively studied decision-making biases, or decision making techniques to counter recognized biases.

uncovered many misconceptions regarding ‘good teaching,’ found systematic causes for deficiencies in teaching, and provided the discovery that a significant number of advanced students, and even physics professors, manage to achieve an impressive academic record without mastering relatively elementary mathematical, and physical science concepts. Arons pioneered new ways to teach math and science concepts based on this extensive research[12].

Here then are ten critical thinking process deemed foundational for decision making activities. Examples are provided to elucidate the meaning of each process:

1. **Consciously raising the questions “What do we know…? How do we know…? Why do we accept or believe…? What is the evidence for…?” when studying some body of material or approaching a problem.**

Consider the assertion, which most people will make, that the moon shines by reflected sunlight. How many people are able to describe the evidence, available to anyone who can see, that leads to this conclusion? Esoteric intellectual skills are not required, even young children can follow and understand. All that is required is to lead them to observe the locations of both the sun and the moon as a few days go by. Yet for the vast majority of the population the “fact” that the moon shines by reflected sunlight is received knowledge, not sustained by understanding.

Similar questions should be asked and answered in other disciplines: How does the historian come to discover how the Babylonians lived? What is the evidence that supports the claim that particular monetary policies promote economic stability? What is the evidence that the structure of matter is discrete rather than continuous? What do we mean by the terms oxygen and nitrogen, and how do we recognize these as different substances?

Cognitive scientists have identified two major classes of knowledge: figurative or declarative (involving semantic memory) and then operative or procedural (involving episodic memory). Declarative knowledge or semantic memory consists of knowing facts (matter is composed of atoms and molecules; terrorists attacked the United States on September 11, 2001; Grandma Jones likes pumpkin pie). Operative knowledge or

---

[12] Dr. Arons earned M.E., M.S. and Ph.D. degrees in physical chemistry (Stevens Institute of Technology, 1937 and 1940; Harvard 1943), and then worked at Woods Hole Oceanographic Institute (WHOI) during and shortly after WWII. His work included leading the team that measured the shock waves of the first atomic bomb tests at Bikini Atoll in 1946. Upon leaving WHOI, Arons taught at Stevens Institute of Technology, then Amherst College, and then spent the remainder of his teaching career at the University of Washington. Dr. Arons was a Fellow of the American Physical Society and the American Association for the Advancement of Science, and was a member of the American Geophysical Union, National Science Teachers Association, and the American Association of Physics Teachers (AAPT), which he served as President in 1967. He was the 1972 recipient of AAPT's Oersted Medal, given in recognition of his notable contributions to the teaching of physics. While at Amherst College he was widely know as a skilled teacher and was featured in a cover story on education in Time magazine. “The content of the course for which he was cited and the instructional philosophy on which it was based are exemplified in his text, *Development of Concepts of Physics*, noted the AAPT Oersted Medal citation. The very careful attention of a logical sequencing of ideas, the deep concern for a careful development of concepts in the minds of students, and the steady attention to the cultural basis of Western science so evident in that book have been his hallmarks. Arons also wrote other texts, including *A Guide to Introductory Physics Teaching*, which reflected his interest in training teachers and non-science majors” (WHOI 2001).

episodic memory involves understanding where the declarative knowledge comes from or what underlies it (Anderson 1980; Kandel, Kupfermann et al. 2000). Operative knowledge also involves the capacity to use, apply, transform, or recognize the relevance of declarative knowledge in novel situations.

Preschool children nearly always ask “How do we know? Why do we believe? Why? ...Why?” questions until formal education teaches them not to do so. Arons has observed that most high school and college students must be pushed, pulled, and cajoled into stating and examining such questions; they do not do so spontaneously. Arons (1990) continues:

> …our usual pace of assignments and methods of testing all too frequently drive students into memorizing end results, rendering each development inert. Yet given time and encouragement, the habit of inquiry can be cultivated, the skill enhanced, and the satisfaction of understanding conveyed. The effect would be far more pronounced and development far more rapid if this demand were made deliberately and simultaneously in science, humanities, history, and social science courses rather than being left to occur sporadically, if at all, in one course or discipline (p. 315).

2. **Being clearly and explicitly aware of gaps in available information. Recognizing when a conclusion is reached or a decision made in absence of complete information and being able to tolerate the ambiguity and uncertainty. Recognizing when one is taking something on faith without having examined the “How do we know…? Why do we believe…?” questions.**

Fascinating investigations of cognitive skill and maturity are conducted by administering homework problems or test questions in which some necessary piece of information has been deliberately omitted, and the question cannot be answered without getting the information or making a plausible assumption to fill the gap. “Most students and many mature adults perform very feebly on these tests. They have had little practice in such analytical thinking and fail to recognize, on their own, that information is missing. If they are told that this is the case, some will identify the gap on reexamining the problem, but many will still fail to make the specific identification” (Arons 1990, p. 315).

How do such deficiencies emerge and persist? One may recall that it is not always inappropriate to ask people to take something on faith. This is occasionally necessary when developing a concept, but this should never be done without making the participants explicitly aware of what evidence is lacking and exactly what is being taken on faith. Failing to do this will prevent people from establishing a frame of reference from which to judge their level of knowledge. They will subsequently lose the ability to discriminate clearly those instances where evidence has been provided from those in which it has not.

One area in which the author has observed decision makers showing considerable deficits in recognizing when something is being taken on faith (i.e., without carefully addressing the “How do we know…? Why do we believe…?” questions) involves the critical assumptions of independence and identical distribution in quantitative risk assessments of complex systems such as aircraft, spacecraft, and nuclear power plants.

For example, recall the experiment involving tossing a fair coin (e.g., a U.S. penny) where one is concerned with the number of times the head of the penny is showing after the toss. It is reasonable to assume that each time the penny is thrown there is a 50/50 chance that a 'head' will result, i.e., 1 out of 2 possible outcomes or ½ is the likelihood that after one toss you will have observed one 'head' and (½)(½) = ¼ is the likelihood that after two coin tosses you will have observed two 'heads.' Now are sequential coin tosses truly independent? Or, for that matter, is one guaranteed that the coin toss will result in a 'head' or a 'tail'? [13]

The assumption of independence of events is very difficult to prove when complex systems are involved. This is so for two main categories of reasons: the first is that one cannot logically conceive of all of the possible dependencies that can be present in a complex system (recall the unknown, unknowns). The second is that observational data is always very limited regarding complex systems due to resource limitations (e.g., time, money). Even in the case where a great deal of time and money is available, as the number of variables in the system alleged to be important increases, the number of observations required for confident understanding of the system increases exponentially. In other words, the more detailed you wish your system understanding to be, the more difficult it becomes to collect the data required for placing system events on the continuum ranging from completely independent to completely dependent.

A close neighbor to this issue of dependency is one of system stability: that is, each time you observe the system 'repeating' performance of a particular event, is it actually repeating the event or does the system actually contain important differences each time the 'experiment' is run and observations are recorded? Are the events indistinguishable from one another (i.e., identical)? We know that, "…the very act of repeating an experiment under identical conditions an infinite number of times is physically impossible" (Larsen and Marx 2001). But are the system events under study 'close enough' to the desired modeling assumptions? Complex systems rarely, if ever produce independent, identically distributed events like those studied in probability theory and statistics courses. See Ross (2000) for a complete definition of independence and a discussion of independent identically distributed events.

The purpose of mentioning the limitations of applying probability theory here is not to cause despair over the use of such methods. The author has been extensively trained in probability theory and statistics, recognizes the power of these tools, and applies those tools often. Yet it is disappointing to find risk analysts who have forgotten the great number of simplifying assumptions they have made (often implicitly) when analyzing very complex systems, touting with *unqualified confidence* that an event has a likelihood of 1 chance in 100,000,000,000 or even less.

---

[13] As an interesting aside, the author once threw a handful of small change onto a table top while clearing his pant pockets at the end of the day. Amazingly, a U.S. penny landed on its edge, rolled approximately 1 cm and remained standing! Therefore, the toss did not reveal a tail or a head, but only the thin edge of the penny was facing up from the table surface. This is truly an astonishing occurrence when one examines the dimensions and typical imperfections from symmetry of any particular U.S. penny (not to mention table top aberrations). The lesson: many occurrences thought to be incredibly unlikely actually happen–in high consequence systems as well–often this is so due to unforeseen dependencies, sometimes it is due to random variations.

This type of analyst has forgotten that such quantitative analyses for very rare events in complex systems (i.e., where little if any observational data are available) are mainly of benefit in judging relative risks of competing alternatives in system design and for prioritizing (on an ordinal scale[14]) which system component vulnerabilities should be addressed first. Furthermore, the discipline and structured investigation imposed by applying a quantitative, reductionist analysis of a complex system does improve knowledge of the strengths and weakness of a complex system (Knief 1992). Notwithstanding such benefits of risk analysis, aleatory and epistemic uncertainties inherent in the system representation should inspire analysts to be eternally humble and filled with a healthy skepticism regarding the robustness of their analyses. The nuclear power industry guards against the uncertainty of undesirable events (known unknowns and unknown unknowns) by employing an elaborate strategy referred to as defense-in-depth as will be discussed in section seven of this report.

3. Discriminating between observation and inference, between established fact and subsequent conjecture.

Many people have a severely impoverished ability to make such discriminations. They are not accustomed to keeping track of logical sequence, and they are frequently confused by technical jargon they may have previously been exposed to, but never clearly understood. For example, in the case of the illumination of the moon introduced under the first critical thinking process, students must be made explicitly conscious of the fact that they *see* the extent of illumination growing steadily as the angular separation of the moon and sun increases with full illumination achieved at a separation of 180°. This direct observation leads to the *inference* that what is observed is actually reflected sunlight. Here is another interesting example from Arons (1990):

> In working up to the concept of "oxygen" (without any prior mention of this term at all) with a group of elementary school teachers some years ago, I had them do an experiment in which they heated red, metallic copper in an open crucible and weighed the crucible periodically. What they saw happening, of course, was the copper turning black and the weight of crucible and contents steadily increasing. When I walked around the laboratory and asked what they had *observed* so far, many answered, "We observed oxygen combining with the copper." When I quizzically inquired whether that was what they had actually *seen* happening, their reaction was one of puzzlement. It took a sequence of Socratic questioning to lead them to state what they had actually seen and to discern the *inference* that something from the air must be joining the copper to make the increasing amount of black material in the crucible. It had to be brought out explicitly that this "something from the air" was the substance to which we would eventually give the name "oxygen." What they wanted to do was to use the technical jargon they had acquired previously without having formed an awareness of what justified it (p. 316).

[14] For an excellent discussion of the nominal, ordinal, interval, and ratio scales of measurement, refer to Stevens (1946).

4. Recognizing that words are symbols for ideas and not the ideas themselves. Recognizing the necessity of using only words of prior definition, rooted in shared experience, in forming a new definition and in avoiding being misled by technical jargon.

From the didactic manner in which concepts are hurled at students in early schooling, it is not surprising that they acquire virtually no sense of the process of operational definition. They come to view concepts as rigid, unchanging entities with one absolute significance that the erudite automatically "know" and that the breathless student must acquire in one intuitive gulp.

> It comes as a revelation and a profound relief to many students when they are allowed to see that concepts evolve; that they go through a sequence of redefinition, sharpening, and refinement; that one starts at crude, initial, intuitive levels and, profiting from insights gained in successive applications, develops the concept to final sophistication (Arons 1990, p. 316).

For example, let us consider a semantic problem that originates for many at an early age stemming from questions such as, "Why do things fall down?" A common answer to the child's question is, "Because of gravity." Arons (1990) has observed that most college students would answer in the same manner. Children, and many adults, take such an answer literally; that is, since the word "because" was invoked, they uncritically jump to the conclusion that a reason has been provided–the "why" is thought to have been answered. They naively believe that a scientific name or technical term provides a reason (Arons 1990).

First, to explicitly alert children or adults to this problem, it may help to explore the history of the term. The Greeks endowed objects with the teleological[15] properties of "gravity" and "levity," which represented built-in tendencies to seek the center of the earth or to rise toward the celestial sphere; seventeenth-century science abandoned both the teleology and "levity" and supplied the name "gravity" to the observed interaction between earth and other objects. Newton's term is attached to the realization, how ever it might work, that the same effect that persuades an apple to fall also binds the moon to the earth, and the planets to the sun (Arons 1990).

Second, following the brief history lesson, it should be explicitly stated that the name (in this case "gravity") does nothing more than conceal ignorance; that is, notwithstanding the beauty of Newtonian Mechanics and the General Theory of Relativity, we have no mechanism for the interaction and no understanding of how it "works." The reader might now be inclined to ask, "What constitutes a 'mechanism,' or explanation of how it works?" Arons responds by stating that "mechanism" or "works" refers to processes that one can visualize in terms of ordinary sense experience. Microscopic effects such as gas pressure and diffusion, evaporation and condensation of liquids, and crystallization and structure of solids, can be visualized in terms of familiar macroscopic particles. Elastic waves in solids may be visualized in terms of what can be seen on soft springs and

[15] Teleology–the fact or character attributed to nature or natural processes of being directed toward an end or shaped by a purpose (Mish 1996).

propagation of electromagnetic waves may be visualized in analogous terms as mechanical shear waves. Corresponding avenues of visualization are not available for quantum mechanical effects or gravitational interactions, "virtual" entities notwithstanding (Arons 1990).

The benefit to realizing that names, as such, do not constitute knowledge or understanding, and combining this fact with operational[16] definitions of terms helps one to recognize when they do not know the meaning of a technical term and to recognize when meaning has, or has not been given. It is unfortunate when one is, "…misled into thinking that they have pursued inquiry and possess understanding when they have, in fact, only used technical terms the meaning of which they did not comprehend and only dealt in vacuous generalizations devoid of genuine thought and substance" (Arons 1990, p. 302). One is wise to remember that words are symbols for ideas and not the ideas themselves.

5. Probing for assumptions (particularly the implicit, unarticulated assumptions) behind a line of reasoning.

In the sciences this tends to be a relatively easy exercise. Idealizations, approximations, and simplifications are not hidden; quite often they are clearly articulated. They are still overlooked by students, however, because explicit recognition and restatement are rarely, if ever, called for on tests or examinations. In history, humanities, and the social sciences (key elements of a classical or liberal education), underlying assumptions are frequently more subtle and less clearly articulated; probing for them requires careful and self-conscious attention among both instructors and students.

The following hypothetical example reveals the importance of probing for underlying assumptions: it was 1976, and a large group of college students were enjoying the energized atmosphere of an anti-nuclear power rally held on the campus. It was an early fall evening, the sun having just disappeared below the horizon, a deep red glow reflected from the scattered clouds above to illuminate the grassy fields, large oak trees, and stately university buildings surrounded this gathering of concerned young people– this picturesque scene was less than 20 miles from the nearest nuclear power plant. After the gentle fade-out of *Mercy Mercy Me (The Ecology)* by Marvin Gaye[17] over the loud speakers, a young man began speaking over the microphone at the makeshift podium and proceeded to regale his peers with the dangers of nuclear power. At one point in his diatribe he declared, "Nuclear power isn't any good for generating electricity–those plants are only 33% efficient! Can you believe it?" The crowd replied with a crescendo of "No!" and gasps of surprise could be heard. The speaker continued, "Yes, as shocking as it sounds, they're only 33% efficient!"

Now let's focus on the 'thoughts' of three different students in attendance at this rally.

---

[16] Recall that operational knowledge involves understanding where the declarative knowledge comes from or what underlies it, and also involves the capacity to use, apply, transform, or recognize the relevance of declarative knowledge in novel situations.
[17] The second single off the popular Marvin Gaye album titled: *What's Going On* that was released in 1971.

The first was studying to be an elementary school teacher and was just a semester away from graduation. She ‘knew’ that outstanding performance for any system is at or near 100% of its potential; the thought of a nuclear power plant being just 33% efficient sounded scandalous to her. The second student was a sophomore industrial engineering student who had some introduction to thermodynamics and vaguely remembered that energy efficiencies of any man-made system tend to be pretty low. But he knew ‘Jimmy’ (the speaker belting out these recriminations against nuclear power) and assumed that 33% referred to the performance of these plants relative to the theoretical potential that they have for converting energy from the fission process to electricity, thus the plants should at least be in the range of 70% or more of that potential. The third student was a mechanical engineering student with a semester to go before walking the stage. He was very familiar with thermodynamics and the efficiencies of engines, motors, and various chemical conversion processes. Being a research assistant for the past year, he’s had to review that knowledge and apply it on several projects. Knowing that the conversion of an energy source into useful work (e.g., mechanical or electrical) often ranges from less than 10 % to 40% or so–especially for complex systems like a power plant which transfers energy from a heat source to steam generation to turbine rotation to electric current generation–he was immediately suspicious.

Student number three suspected that someone had given ‘Jimmy’ some thermodynamic energy conversion number and passed it off as a bad thing. Jimmy never mentioned any assumptions behind the 33% number, such as what practical limitations there may be for the water-steam temperature differentials in the plant. Nor did he mention the differences in fuel costs between nuclear power plants and other plants such as coal or natural gas–why weren’t economic efficiency numbers specifically stated? Only an ill-defined number of 33% efficiency was declared. This student no longer felt comfortable attending this anti-nuclear rally as he suspected that the truth regarding nuclear power was being dramatically distorted. Student number three left the rally.

The above example highlights different levels of omissions in probing for assumptions for each of the three students. The first student implicitly assumed that 100% was the attainable ideal performance level. For her, the possibility that 100% would be an impossible goal using energy analysis norms may have been an ‘unknown unknown.’ That is, she would not have realized that implicit assumption should be called into question. The second student had some engineering training, and should have sought the analysis assumptions, but informal bonds of friendship led him to immediately interpret the situation in a way that ‘made sense’ given his current knowledge base; this neglect of thoroughly probing for data assumptions had the two-fold benefit of avoiding conflict with his friend and reinforcing the norms of the anti-nuclear student body in attendance. The third student appears to have had the richest base of knowledge and experience to decipher the efficiency claims and be suspicious of the claimed ‘inefficiency.’ Unfortunately, the situation did not lend itself easily to questioning the speaker directly about the underlying assumptions and we do not know if this student eventually sought additional data describing nuclear power plant performance levels. Therefore, to greater or lesser degree all three students evidenced important omissions with regard to probing the assumptions of the speaker–so we never discover what 33% efficient really meant in this case.

6. **Drawing inferences from data, observations, or other evidence and recognizing when firm inferences cannot be drawn. This subsumes a number of processes such as elementary syllogistic reasoning (e.g., dealing with basic propositional, "if…then" statements), correlational reasoning, recognizing when relevant variables have or have not been controlled.**

Distinct from the analysis of another's line of reasoning is the formulation of one's own. "If…then" reasoning from data or information must be accomplished without prompting from an external "authoritative source." One must be able to discern possible cause-and-effect relationships in the face of statistical scatter and uncertainty (recall both aleatory and epistemic uncertainty). One must be aware that failure to control a significant variable prevents the possibility of inferring a cause-and-effect relation. One must be capable of discerning when two alternative models, explanations, or interpretations are equally valid and cannot be discriminated solely on logical grounds. Arons (1990) provides us with an example:

> …I present a case I encounter very frequently in my own teaching. When students in a general education science course begin to respond to assignments leading them to watch events in the sky (diurnal changes in rising, setting, and elevation of the sun, waxing and waning of the moon, behavior of the stars and readily visible planets), they immediately expect these naked eye observations to allow them to "see" the "truth" they have received from authority, namely that the earth and planets revolve around the sun. When they first confront the fact that both the geo- and heliocentric models rationalize the observations equally well and that it is impossible to eliminate one in favor of the other on logical grounds at this level of observation, they are quite incredulous. They are shocked by the realization that either model might be selected provisionally on the basis of convenience, or of aesthetic or religious predilection. In their past experience, there has always been a pat answer. They have never been led to stand back and recognize that one must sometimes defer, either temporarily or permanently, to unresolvable alternatives. They have never had to wait patiently until sufficient information and evidence were accumulated to develop an answer to an important question; the answer has always been asserted (for the sake of "closure") whether the evidence was at hand or not, and the ability to discriminate decidability versus undecidability has never evolved (p. 317).

This type of example, where explicit guidance is provided, can play a powerful role in opening people's minds to spontaneous assessment of what they know and what they do not know, of what can be inferred at a given point in time and what cannot.

7. **Performing hypothetico-deductive reasoning; that is, given a particular situation, applying relevant knowledge of principles and constraints and visualizing, in the abstract, the plausible outcomes that might result from various changes one can imagine to be imposed on the system.**

Opportunities for such thinking abound in nearly every educational course. Yet students are often given very circumscribed questions that do not open the door to imaginative hypothetico-deductive reasoning. These limited situations are important and provide necessary exercises as starting points, but they should be accompanied by questions that motivate the student to invent possible changes and purse the plausible consequences.[18]

8. Discriminating between inductive and deductive reasoning; that is, being aware when an argument is being made from the particular to the general or from the general to the particular.

The concepts of “electrical circuit,” “electric current,” and “resistance” can be induced from simple observations made with electric batteries and arrangements of flashlight bulbs. This can then be used to construct a “model” of how an electrical circuit operates. The model then forms the basis for deductions, that is, predictions of the brightness of bulbs in new configurations or when something anomalous (e.g., short circuiting) is imposed on an existing configuration. This type of hypothetico-deductive reasoning can be applied in any domain (e.g., economic models, energy use models, etc.), but one should always be conscious of the distinction between the inductive and the deductive modes and the evidence justifying a particular mode at a particular time.

An additional example highlights how a person may subtlety transition from induction to deduction, possibly without conscious recognition of the act: consider an intelligence analyst who performs incremental analysis, that is, they have formed some initial hypothesis of what data sources may be important and then they regularly receive small bits of data which they may then associate or link to the initial hypothesis. The initial hypothesis can be seen as a first step toward deduction, but overall this should really be seen as an inductive, exploratory process for some time. For example, as individual satellite images arrive, they may be forced to fit into a bigger picture alluded to by past information sources organized from very weak initial hypotheses. This effect may quickly be magnified by organizational pressures favoring consistent interpretations particularly when a previous hypothesis has been written down/submitted in an earlier brief or document (Heuer 1982). This sequence of bypassing appropriate critical thinking processes regarding inductive and deductive reasoning is common and will be discussed in detail later in this report under the description of ‘confirmation biases.’

9. Testing one’s own line of reasoning and conclusions for internal consistency and thus developing intellectual self-reliance.

---

[18] The author has observed, on more than one occasion, instructors who intentionally restrict homework and test questions to very circumscribed cases (ones which students find easy/comforting), because the instructor does not believe that students could really be expected to learn important subtleties of the material in the time allotted, or simply to keep the students happy–this may be exacerbated in the increasingly familiar situation where student’s ratings of teacher performance have considerable influence on teacher’s raises, promotions, etc. Arons has found that such practices seriously undermine the development of critical thinking processes. This may prove to be a negative side effect of so-called ‘grade inflation’ practices. Specifically, the type of grade inflation in which course material is simplified so that achieving high grades becomes relatively easy; hence a large number of students get high grades and leave their courses with an inflated perception of their intellectual abilities in the subject(s) covered. Arons argued that it is better to cover less of the subject matter, but to do so in a manner that hones students’ hypothetico-deductive capacities. The currently popular mantra for creative thinking, i.e., ‘thinking outside of the box’ is closely related to this critical thinking process.

In a technologically advanced society, it is certainly not possible to teach people all they will need to know in a formal education setting. "The principal function of education–and higher education in particular–must be to help individuals to their own intellectual feet: to give them conceptual starting points and an awareness of what it means to learn and understand something so that they can continue to read, study, and learn as need and opportunity arise, without perpetual formal instruction" (Arons 1990, p. 318). In order to learn on one's own (not merely accumulating facts) requires the capacity to assess when understanding has been achieved and to draw conclusions and make inferences from acquired knowledge. Inferring, involves testing one's own thinking process, and the results of that process for correctness, or at a minimum for coherence and consistency. This is a highly sophisticated level of intellectual activity, and to achieve competency one must first be made aware of such processes and their importance then be provided with practice and guidance.

In science courses, they should be required to test and verify results and conclusions by checking that the results make sense in extreme or special cases that can be reasoned out simply and directly. They should be led to solve a problem in alternative ways when that is possible. Such thinking should be conducted in both quantitative and qualitative situations. In the humanities and social sciences, the checks for internal consistency are more subtle, but they are equally important and should be cultivated explicitly. Students should be helped to sense when they can be confident of the soundness, consistency, or plausibility of their own reasoning so that they can consciously dispense with the teacher and cease relying on someone else for the "right answer"[19] (Arons 1990).

10. Developing self-consciousness concerning one's own thinking and reasoning process.

This is perhaps the highest and most sophisticated reasoning skill, presupposing the others that have been presented. It involves reflecting upon the processes one is using, deliberately invoking those most appropriate to the given circumstances, and providing the basis for conscious transfer of reasoning methods from familiar to novel contexts. "Given this level of awareness, one can begin to penetrate new situations by asking oneself probing questions and constructing answers. Starting with artificial, idealized, oversimplified versions of the problem, one can gradually penetrate to more realistic and complex versions. In an important sense, this is the mechanism underlying independent research and investigation" (Arons 1990, p. 319).

These critical thinking processes, summarized in Figure 3.1 are proposed as foundational for decisions regarding high-consequence, highly sophisticated technological systems. It is further proposed that these capacities are critical for being able to properly understand and apply the specific information presented throughout this report regarding biases and

[19] The author has a suspicion that widespread acceptance of moral relativism (i.e., the belief that societies and individuals alone decide what is right and what is wrong–there is no 'truth' in such matters aside what me and my neighbor agree upon) may have some spillover effect on more traditionally 'quantitative' matters. Thus 'truth' in the world, in a broad sense, becomes more of a democratically agreed upon construct than one informed by disciplined application of 'critical thinking processes' as defined in this report. For short, entertaining introductions to moral relativism (a.k.a., liberal social philosophy) refer to Fink (1981) and Budziszewski (1999).

the decision-making approach. The reader is encouraged to consult Arons (1990) for his complete discussion of this list of processes, supporting evidence, and specific strategies for nurturing critical thinking skills in the teaching of math and physical science.

Critical thinking processes

1. Raising the **questions**: "What do we know…? How do we know…? Why do we accept or believe…? What is the evidence for…?
2. Clear and explicit awareness of **information gaps** (i.e., recognizing when one is taking something on faith).
3. Discriminating between **observation and inference**, between established fact and subsequent conjecture.
4. Recognizing that **words are symbols for ideas** and not the ideas themselves. Recognizing the necessity of using only words of prior definition, rooted in shared experience, in forming a new definition and in avoiding being misled by technical jargon.
5. **Probing for assumptions** behind a line of reasoning.
6. **Drawing inferences** from data, observations, or other evidence and recognizing when firm inferences cannot be drawn (i.e., **inference adequacy check**).
7. **Hypothetico-deductive reasoning**; apply relevant knowledge of principles and constraints, and abstract visualization of plausible outcomes from imagined changes imposed on the system.
8. Discriminating between **inductive and deductive reasoning**; that is being aware of when an argument is made from the particular to the general or from the general to the particular.
9. Test one's own line of reasoning and conclusions for **internal consistency**.
10. Develop **self-consciousness** concerning one's own thinking and reasoning processes.

Figure 3.1. Summary of critical thinking processes appropriate for supporting decision making in high consequence domains. Adapted from Arons (1990).

4. Decision Making in the Technologically Advanced Society

The achievements of science and technology have been important, impressive, and have helped to provide unprecedented physical safety and material wealth. Yet they have done so at the cost of substantially increasing individuals' vulnerability to risks associated with interdependence (Freudenburg 2001). This section summarizes reflections by Freudenburg (2001), regarding risk-related decision making and, in particular, how the nature of a technologically advanced society as opposed to an agrarian/simple craft-based society amplifies the need for trust between technology experts, high-level decision makers (e.g., politicians), and members of the general public. This section ties strongly to the critical thinking processes introduced in the previous section. Furthermore, this section suggests who ought to be included in high-consequence, technologically advanced decisions and when technology experts or the general citizenry deserve the strongest vote on aspects of those critical decisions.

Freudenburg begins his discussion by reflecting upon public sentiments he heard while being on a U.S. National Academy of Sciences National Research Committee reviewing the analyses conducted for a nuclear waste storage site. Comments from the citizens included the classic, “Scientists have lied to us before, and scientists will lie to us again” (Freudenburg 2001, p. 125). This reflected a sentiment widely held by fellow citizens in attendance that scientists habitually put their organizational interests ahead of even-handed analysis. Freudenburg quickly learned that he was not dealing with a group of extremists who irrationally opposed the nuclear waste dump. What he found were people who had initially supported the idea of a ‘scientific process’ for analyzing proposed sites, but who had become, over the course of the site selection process, convinced that the process was neither fair, nor scientific.

> Although many of us initially believed that the citizens were simply being sore losers, we ultimately learned something else: in fact, the process had included not just technical errors, but many cases in which arbitrary decisions had been characterized as scientific ones. Understandably–at least from the point of view of the proponents–whenever the arbitrariness was questioned, the critics were accused of being ‘anti-science.’ The proponents’ intent of course, had been to improve the credibility of the siting process by wrapping it in a flag of science. The result, however, had been virtually the opposite; what actually took place was instead a serious drop in the credibility of science. The net unintended effect, in other words, had been the creation of a self-fulfilling prophecy–accusing a wide range of people of being anti-science, and then ultimately creating just that result (Freudenburg 2001, p. 125).

Once the technical experts and the general citizenry[20] descend into this polarized relationship, a phenomenon called the ‘spiral of stereotypes’ may emerge in which people on opposing sides of an issue stop talking to one another, but not about one another. That is, rumors and a priori beliefs that are already accepted in each camp take on more credibility than dialogue intended to promote mutual understandings between the camps. The us/them thinking takes over and ‘rationality’ is assumed to characterize the ‘us’ members of the debate, and anyone who shows the poor judgment to disagree with ‘us’ must therefore be exhibiting irrationality (Freudenburg 2001).

Before proposing a solution for avoiding a polarized decision making environment, it may be helpful to explore some of the subtleties that encourage such a polarization. It is proposed that many arguments about the proper roles of experts and citizens in high-consequence, technologically complex decision making emerge due to a fundamental misunderstanding of what it means to live in a ‘technologically advanced’ society. For instance, there have been some broad indications of declining public confidence in

[20] It is very important to emphasize that this section is analyzing the ways in which level-headed citizens become opposed to high consequence, technologically advanced systems and how appropriate inclusion of such citizens is a requirement for ‘appropriately’ making decisions regarding these systems; it does not address extreme groups who have abandoned all hope or interest of achieving ‘mutual understanding.’

science and technology and the experts who defend them. How can this sentiment prevail when dramatic reductions in risk can be attributed to advancing technology? The prime example being the dramatically reduced risk of death, measured by life expectancy data (i.e., more defenses against and treatments for disease and physical injury, improved shelters with heating and cooling, better plumbing and general hygiene practices, improved quantity and quality of food, safer means of transportation, safer means to produce food and durable goods, etc.). To many, such improvements make them feel justified in denouncing many or all public concerns toward science and technology as being ignorant, ill-informed, and/or irrational. At least three key problems with this type of thinking have been identified (Freudenburg 2001).

The first problem is the personal affront caused by telling someone they must be ignorant, ill-informed, and/or irrational. This affront may be indirectly expressed by expending considerable effort to exclude the public from a decision. Either form of insult tends not to engender receptivity of anything else that follows. The second problem is one of adaptation, scientific and technological achievements rapidly become part of people's baseline expectations. There is likely to be very little persuasive power to the argument that all should be happy to live as long as did our great-great-grandparents– especially when this implies that anyone over 40 years of age ought to be very happy about not being dead (Freudenburg 2001).

The third problem is the most serious, most often overlooked and has been termed the 'fundamental misunderstanding' about what it means to live in a technologically advanced society. We begin with a question: Do people living today (i.e., in a technologically advanced society) actually 'know more' than their great-great-grandparents did? Collectively, the answer is a resounding 'yes.' Individually, however, we know much less today than did our great-great-grandparents about the tools and technologies on which we depend (Freudenburg 2001).

> In the early 1800s, roughly 80% of the US population lived on farms, and for the most part, those farm residents were capable of repairing, or even of building from scratch, virtually all of the tools and technologies upon which they depended. By contrast, today's world is so specialized that even a Nobel laureate is likely to have little more than a rudimentary understanding of the tools and technologies that surround us all, from airliners to ignition systems, and from computers to corporate structures (Freudenburg 2001, p. 128).

Therefore, people living in a technologically advanced society tend not so much to be in control of technology as dependent upon technology. Our increases in *technical control* have led in a sense to a decrease in *social control*. Stated simply, the problem is that people now must depend on whole armies of specialists and specialized organizations, most of whom we will never meet, let alone be in a position to control.

The failure of an expert, or specialized organization to accomplish the job required has been defined by some as *recreancy*[21] (Clark and Short 1993; Freudenburg 1993).

[21] The term *recreancy* is used to depersonalize/reduce the intensity of the negative emotional connotations associated with these types of failures. For example, a particular engineer (recognized by many as an

Recreancy, and its close neighbor trustworthiness, have been shown by systematic research to be key factors in the increasingly acrimonious chemistry between technical experts and members of the general public. For example, an analysis of attitudes regarding a proposed low-level nuclear waste facility found that sociodemographic variables were weak predictors of attitudes whereas measures of recreancy were found to be very strong predictors of attitude (Freudenburg 2001).

Thus we have the three problems with dismissing people[22] who express concerns regarding decisions involving high-consequence, advanced technologies: personal affront; increasing expectations for technological advancements; and increasing dependency upon advanced technology (i.e., increasing need for trust). These three problems, combined with the folly of experts who try to pass off arbitrary decisions as scientific decisions, provide a proven recipe for creating a polarized decision-making environment. A polarized decision making environment will severely frustrate application of the bias mitigation insights, critical thinking processes, and decision making approach proposed in this report. So how does one avoid this very large decision-making trap? General suggestions are described below and summarized in Figure 4.1:

1. Do not automatically brand members of the public who raise concerns over decisions involving high consequence, advanced technologies as ignorant, ill-informed, or irrational.
2. Apply the ten critical thinking processes in the previous section to ensure that technical experts are focused upon **factual or technical questions**. Poor command of the critical thinking processes will prevent proper identification of authentically *factual* or *technical* questions.
3. Identify when aspects of a decision (model assumptions, metrics, etc.) involve answers to questions that are value judgments. Examples of such questions include: Is that safe enough? Is that level of uncertainty in the safety estimate acceptable? How clean is clean enough (i.e., environmental restoration)? When it comes to questions of **values**, another word for 'scientist' is 'voter.' Ordinary citizens have just as much legitimacy in deciding questions of values as do scientists and engineers (Freudenburg 2001). Of course, it is essential to note that when answering a value question a prerequisite may involve providing information and training to allow ordinary citizens to profit from analyses surrounding potential costs and benefits of different answers to the question (i.e., tools to strengthen normative skills, ideally supplied by a trusted source–maybe a 'neutral' party to the decision process).
4. Apply the ten critical thinking processes in the previous to aid in identifying **blind spots**. That is, items that have been overlooked. Additional terms for blind spots

---

honest, trustworthy person who struggles to improve the common good) may, despite best efforts, be confounded by various aleatory and epistemic uncertainties arising in a complex system that leads to a 'failure to accomplish the job.' Words such as 'lazy,' 'untrustworthy,' 'incompetent,' 'foolish,' 'stupid,' etc., would not be appropriate in this type of situation.

[22] The reader is alerted to the fact that many times the non-expert 'people,' in a decision involving high-consequence, advanced technologies, also includes high-level policy makers who will necessarily have considerable power in determining final decisions (esp., national security decisions where close scrutiny by the 'general public' is not allowed). This highlights the need for a structured, defensible process which includes both subject matter experts and non-experts at appropriate junctures.

include *epistemic uncertainties* and unknown unknowns, which were discussed earlier in this report. Ordinary citizens are also very helpful for contributing to the discussion of blind spots as their thinking tends to be bounded much differently than those of a particular community of 'experts.' One should keep in mind Perrow's (1984) memorable definition of an *expert* as one who can solve a question much more quickly and efficiently than most, but who runs a higher risk than do other citizens of asking the wrong questions altogether.

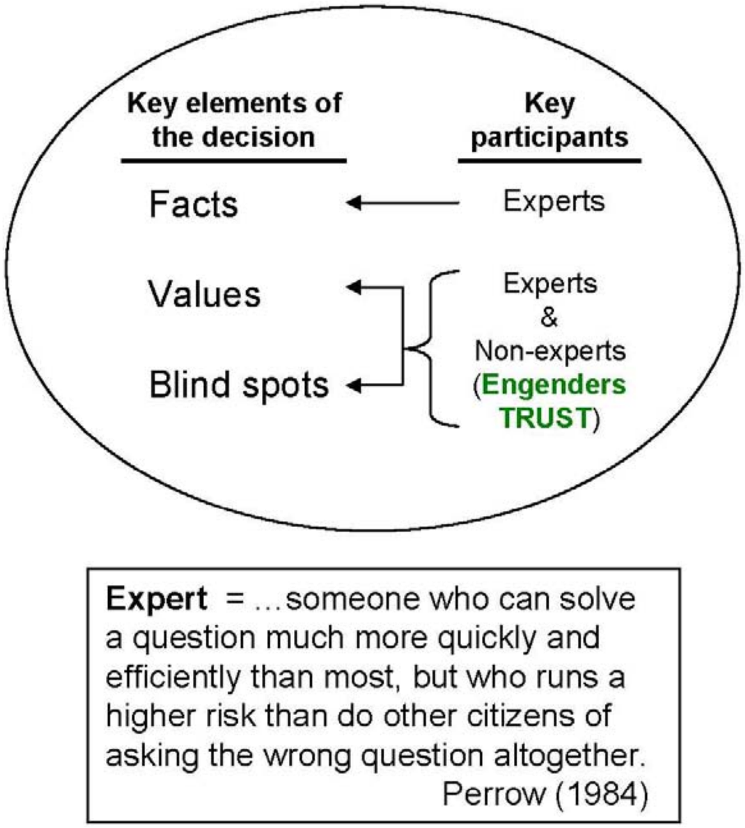

Figure 4.1. Building an appropriate decision making team.

This section regarding decision making in the technologically advanced society provides a simple guide for identifying those members of society who ought to be part of the high-consequence, complex decision process. These members of society should include both experts and non-experts with a prominent role for ordinary citizens who will ultimately participate in (or be subjected to) both the benefits and the costs resulting from such decisions. This advice resonates strongly with recommendations provided by a report from the U.S. Secretary of Energy Advisory Board titled *Earning Public Trust and Confidence: Requisites for Managing Radioactive Wastes* (SEAB 1993). In fact, SEAB (1993) also advises the Department of Energy to consult broadly about the design and implementation of training in *public involvement principles* to aid decision making regarding radioactive waste management. It is envisioned that the contents of this report may prove beneficial to improving those public involvement principles in addition to enhancing such decision making processes in general.

5. Evidence for the Unique Risk Perception and Decision Making Framework

In this section a transition is made from general aspects of decision making, critical thinking processes, advanced technologies and decision team composition, to detailed discussions of risk-related decisions along with perceptual and decision making biases/tendencies uncovered by previous research. Each of the biases/tendencies discussed here are detailed treatments of items briefly introduced in the unique risk perception and decision making framework presented in section two of this report. This section provides a foundation to support the face validity of the unique framework and the associated decision-making improvement strategies.

The human mind has an extraordinary capacity to match many complex patterns and integrate complex information, but this talent seems to be most effective in specific types of situations in which numerous senses are stimulated during complex, task specific cognitive–motor skills such as throwing a baseball or playing a musical instrument, or identifying a dangerous traffic situation while driving along the interstate. This ability has been stated more elegantly by Simon and Sumner (1968), “People appear to have strong propensities, whether innate or learned, to discover patterns in temporal sequence presented by the environment, and to use these evidences of pattern for prediction.” Research in neural science and cognitive psychology has provided strong evidence that many of these tendencies are indeed innate[23] (Gibson and Walk 1960; Fantz 1961; Bower 1966; Kandel, Schwartz et al. 2000). It also appears that people take these types of talents, which are essential to interacting successfully with their immediate environment, and extrapolate them to highly complex phenomenon for which their direct experience is very limited (e.g., forming opinions about the risk of nuclear power from a handful of exaggerated media reports showing frightening graphics of ‘risks’).

Furthermore, while humans are excellent at integrating complex sensory information using hierarchical, distributed, parallel processing features of the central nervous system, it is also important to note how the complexity of our environment or the systems with which we interact still outstrip the analytical capabilities of a single person (Kandel, Schwartz et al. 2000). Given the variety of stimuli to which people can attend in the environment, cognitive processes are highly selective. One result of this need to be selective concerns the fact that people tend to anticipate what they are about to encounter and are more receptive to stimuli that confirm what they are anticipating to encounter (Hogarth 1975). This simplification strategy of tending to expect certain things (i.e., coherence with previous experience or decisions) underlies many of the decision-making biases discussed later in this section.

[23] That is, these tendencies are present at birth. For example, infants are specially tuned to identify human faces and readily encode in memory particular faces (esp. their mother), whereas a disordered or random arrangement of ‘human facial parts’ does not provoke interest. An additional example is that of the infant’s expectation of particular object trajectories when something begins to move. Evidence abounds for innate pattern matching tendencies. In addition to extensive reading of the human development literature, the author & his wife had the pleasure of watching their first-born child develop as this report was being written–it was fascinating to observe her tendencies for attending to certain stimuli and disregarding other similar stimuli–directly in line with the published research on infants.

In high-consequence decision making, the most vexing debates among stakeholders often surrounds potential risks associated with decision options. Therefore, many of the biases/tendencies described here will use risk-related examples. The rational approach to risk management involves maximizing the areas where we have some control over the outcome while minimizing the areas not under our control with the understanding that the linkage between cause and effect is typically hidden from us (Bernstein 1998). In 1738, Daniel Bernoulli (of the famous family of mathematicians) offered his definition of utility; he claimed that different people will pay different amounts for desirable things, and as one accumulates more of that desirable thing, the less they will be willing to pay to acquire more (Bernoulli 1738). Of course, in common experience we see this norm being violated, as Bernstein (1998) has noted, "Today, we recognize that the desire to keep up with the Joneses may lead us to want more and more even when, by any objective standard of measurement, we already have enough" (p. 188). Thus, it may be more accurate for most people that wealth does not motivate your decision, but whether you stand to become richer or poorer (Bernstein 1998).

This is akin to the perception of oneself as a winner or a loser based on the decisions we make (and the resulting consequences) and brings to mind the popular *Law of Effect* which states that people strongly avoid negative stimuli (i.e., pain, discomfort, embarrassment) and seek to increase positive stimuli, i.e., pleasure (Miller 1962; Kandel, Kupfermann et al. 2000). This desire to avoid negative stimuli was illustrated recently by research showing a stronger tendency to believe negative information regarding potential health dangers than positive information. This was true regardless of the credibility of the source of the information (Siegrist and Cvetkovich 2001). One way of interpreting this result is that the 'cost' of being wrong about negative effects is perceived to be more important than the 'benefit' of being wrong about positive effects; i.e., err on the side of caution.

Among the factors that impact decision making are learned behaviors. For instance, the language habits of Western culture tend to promote confusion between certainty and belief. People are encouraged to act and speak as though they are certain of things even when they are only fairly certain. The other end of this phenomenon is acting and speaking as if opinions are worthless when they are only perceived as weak (Savage 1971; Hogarth 1975). These tendencies may assist with simplifying our interaction with those around us, but they are simultaneously discounting the importance of certainty in our reflection about a topic.

People have also been found to be much more likely to accept a higher voluntary risk than a higher involuntary risk (Starr 1969). For example, people tend to be more willing to accept the high risks of driving a car (according to Mokdad, Marks et al. (2004), there are approximately 43,000 motor vehicle accident deaths annually) than to accept high risks as a passenger on an airplane. U.S. airline fatality rates per million flights flown from 1970 through 2004 range from 0 to 5.88 depending on the airline; the average is 0.78[24]. In the first case, the person is controlling the car and may feel more in control of whether they encounter or can avoid hazards; in the second case–the airplane passenger has no control over the actions or inaction of the pilots in the cockpit and must trust that

[24] There were 21 fatalities due to the only U.S. commercial airline crash in 2003 (Airsafe.com 2004).

all will go well. Indeed, the capacity to trust oneself and others is a critical determinant of most decisions.

Translation of this *locus of control*[25] phenomenon to the energy domain can be made readily when one considers the involuntary relationship that many people have with energy production (Wickens and Hollands 2000). An individual does have some choice over whether they live near a nuclear plant or in a city with air pollution aggravated by fossil-fuel burning (yet this decision may also be somewhat involuntary given employment opportunities and family ties). However, they will not have direct control over the operators of those facilities and must trust that safe practices are being observed (Inhaber 2004). Furthermore, most members of the general public do not understand the technical aspects of nuclear power plant operation regardless of where they are living. Recall that individuals become increasingly dependent upon their advanced technologies (i.e., increasing need to trust others regarding matters they do not understand).

In addition to control, it may also be true that the perceived nature of the injury or death is an important contributing factor to the perception of voluntary and involuntary risks. For example, if a person believes that death in an airplane crash will be swift (i.e., little or no suffering) and that death due to a high level of radiation exposure from a nuclear power plant accident will involve days, weeks, or months of agonizing pain–there may be a strong preference for the low-likelihood airplane crash over the extremely low-likelihood radiation exposure from a nuclear power reactor accident. The *perceived nature of the death* or *dread* has been found as an important factor in risk perception by a number of researchers (Geertz 1973; Slovic, Fischhoff et al. 1981; Slovic 1986; Slovic 1987; Jasanoff 1998; Cha 2000; Slovic 2000; Cha 2004). Another aspect of *dread* that plays a significant role in risk perception involves latent toxicity or genetic effects (Fischoff, Slovic et al. 1978). Examples include severe mental impairments in offspring due to mercury exposure (a byproduct of coal burning), or lung cancer due to radon releases (i.e., releases of radioactive radon from coal burning) and a range of genetic anomalies due to radiation exposure (a potential hazard from a nuclear reactor accident).

It may be instructive to explore further the public perception of ‘power plant operation.’ To begin, consider the following question: do most members of the general public understand the fossil fuel burning process for electricity generation any better than the nuclear process for electricity generation? The author suspects that the answer is no with respect to start-to-finish technical understanding of how water is heated to steam, then steam drives rotation of a turbine that subsequently rotates an arrangement of conductive wire through an electromagnetic field to produce the flow of electric current. However, it is likely true that most members of the general public perceive that they understand fossil fuel ‘burning’ much better than nuclear fission ‘burning.’ Many individuals have acquired first hand sensory experience using natural gas stoves, natural gas heaters, charcoal grills, gasoline burning engines and campfires. Therefore they have some

[25] Locus of control was defined by Rotter (1966) as a person’s perception of the control they have over job performance and work-related rewards such as pay and promotion. People identified as having an internal locus of control believe that such things are under their control. Those identified as having an external locus of control believe such things are the result of luck, chance, or whether the boss likes them–i.e., not within their control.

intuitive feel for certain risks (e.g., getting burned, starting an uncontrolled fire, smoke inhalation, etc.), but also feel comfortable with mitigation approaches (e.g., stay a moderate distance away from the heat, put water on the fire, avoid the smoke, turn off the fuel flow, wait a period of time for the fire to burn itself out, etc.). They understand how a problem with the 'burning' may be addressed and may feel confident that they could identify such a problem (e.g., explosion, bright flames, smoke plumes) and then avoid the consequences (e.g., move away from the heat source, seek fresh air).

Such intimate sensory experience is not available for the nuclear 'burning' process; therefore, it is logical that this would decrease confidence in being able to identify a problem (e.g., a high energy radiation source wafting through the air–measurements with specialized equipment are required for this identification), and then making a choice to avoid the consequences (e.g., moving away from the radiation source, covering it, moving behind a shield, taking medications that reduce the occurrence of chemical reactions between particular radionuclides and biological tissues, medications that flush radionuclides out of the body, etc.). This dearth of sensory access to nuclear 'burning' heightens the importance of 'trusting' those in control of the process–more so than in the case of fossil fuel burning.

Interestingly, in the fossil fuel burning case, past sensory experience may prompt an excessive focus on short timescale events such as getting burned by heat and flame or overwhelmed by smoke instead of long timescale events such as life-shortening due to prolonged breathing of polluted air.[26] With nuclear fission, the lack of sensory experience with short timescale events offers no distraction from the potential of cancer in the future due to exposures today. It may also be true that the 'invisible' nature of radiation may prompt people to focus excessively on the possibility of being right next to a powerful radiation source that causes death rapidly[27] or that causes severe genetic defects in future offspring. This 'invisible' nature of radiation is much different than other 'invisible' phenomena such as electricity. Although one cannot 'see' a harmful voltage source (ignoring the hum of transformers, or the 'hair raising' that may portend a lighting strike), they will certainly know right away if they come in contact with such a source–the effects are immediate, and attention grabbing. Furthermore, they usually require very close proximity to the physical source (e.g., outlets, generators, power lines, etc.).

The previously mentioned phenomena of insatiable acquisition (i.e., keeping up with the Jonses, rising expectations, etc.) along with the law of effect, may provide some insight on the drive by some to demand higher and higher levels of safety (e.g., reduced chances of radiation releases from NPPs) far beyond any objective reflection on the true dangers faced by miniscule radiation releases. Of course, there is certainly much more at work behind such radiation-averse behavior. Two of the most respected and influential researchers in the realm of risk and uncertainty are Daniel Kahneman and Amos

[26] This 'polluted air' may or may not be easily identifiable as polluted–not entirely unlike many radiation sources invisible to the human senses.
[27] Rapid death due to radiation is akin to the clinical definition of an acute whole body dose of ~ 300 rem, which is expected to cause death in 50% of those exposed within 30 days if no medical treatment is given (Glasstone and Sesonske 1963).

Tversky, together they have studied human decision making for over 30 years. In organizing their results they have developed what they call Prospect Theory (Kahneman and Tversky 1979; Tversky and Kahneman 1992). In prospect theory, two human shortcomings are emphasized. The first is *emotional responses* that prevent the appropriate level of self-control that is required for rational decision making. The second is that people often *do not understand the domain* (i.e., facts and technical items) at issue. Recently, Tversky and Koehler (1994) have also developed Support Theory which states that the judged probability of an event depends on the explicitness of its description. These decision making 'theories' have emerged from extensive, systematic study of decision making biases.

In the literature regarding probabilistic judgments of events, there are two important types of knowledge that seem to determine how well people perform: (1) *substantive* knowledge is that which regards the topic being explored (e.g., weather forecasting estimates by a meteorologist, stock market analysis by an investment banker, disease incidence in the population by a medical epidemiologist, radiation risks to the public by a health physicist, etc.), and (2) *normative* knowledge is the ability of the person to express their opinions in probabilistic form, e.g., someone with a Ph.D. in statistics would have extensive normative knowledge (Hogarth 1975). The degree to which people possess (or lack) substantive and/or normative knowledge, and the ability to relate those two using *critical thinking processes* (discussed in section 3 of this report) has a large impact on the degree to which the biases below affect decision making skill. We now consider in some detail some of those biases or tendencies related to risk perception and decision making.

First, people have been shown to be *insensitive to sample size*, Tversky and Kahneman (1974) provide a great example of this effect by presenting the following scenario and question: a certain town is served by two hospitals. In the larger hospital about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50 percent of all babies are boys. However, the exact percentage varies from day to day. Sometimes it may be higher than 50 percent, sometimes lower. For a period of 1 year, each hospital recorded the days on which more than 60 percent of the babies born were boys. Which hospital do you think recorded more such days?

- The larger hospital (21 people selected this answer)
- The smaller hospital (21 people selected this answer)
- About the same (that is, within 5 percent of each other) (53 selected this answer)

The correct answer is that the smaller hospital should record more days in which > 60 percent of the babies born were boys. The larger hospital would report results according to the *law of large numbers*[28] which states that the sample mean will converge on the true population mean as the sample size increases (Hines and Montgomery 1990; Larsen and Marx 2001). That is, the larger hospital, with more births recorded, will exhibit a male birth percentage that is closer to the 'true mean'; this is especially true when statistics are gathered over a long period of time (in this case for 365 days).

[28] Some statisticians will refer to this as the *weak law of large numbers* to remind people that this is merely a robust heuristic–not an absolute guarantee for a specific collection of population statistics.

Second, people are fairly good at guessing values of central tendency, such as averages, but they are poor at assessing the variance of data. Instead of thinking in terms of variance (i.e., the square of the standard deviation), it appears that people tend to think in terms of the *coefficient of variation* (i.e., the standard deviation divided by the mean). This is another instance of human perception being highly context dependent. For example, when a person sees the tops of trees in a forest they will often imagine that the tree tops form a relatively smooth surface, even though the trees may vary in height by 10 feet or more. Now if the same person imagines their cluttered desk top at work with a blanket thrown on top, they would likely perceive that as a bumpy and variable surface. The tree tops in the forest are much more variable than the surface of your desk in absolute units of height, but the desk surface is likely more variable relative to the size of the objects under consideration (Peterson and Beach 1967; Hogarth 1975). Figure 5.1 provides a visual illustration of this effect.

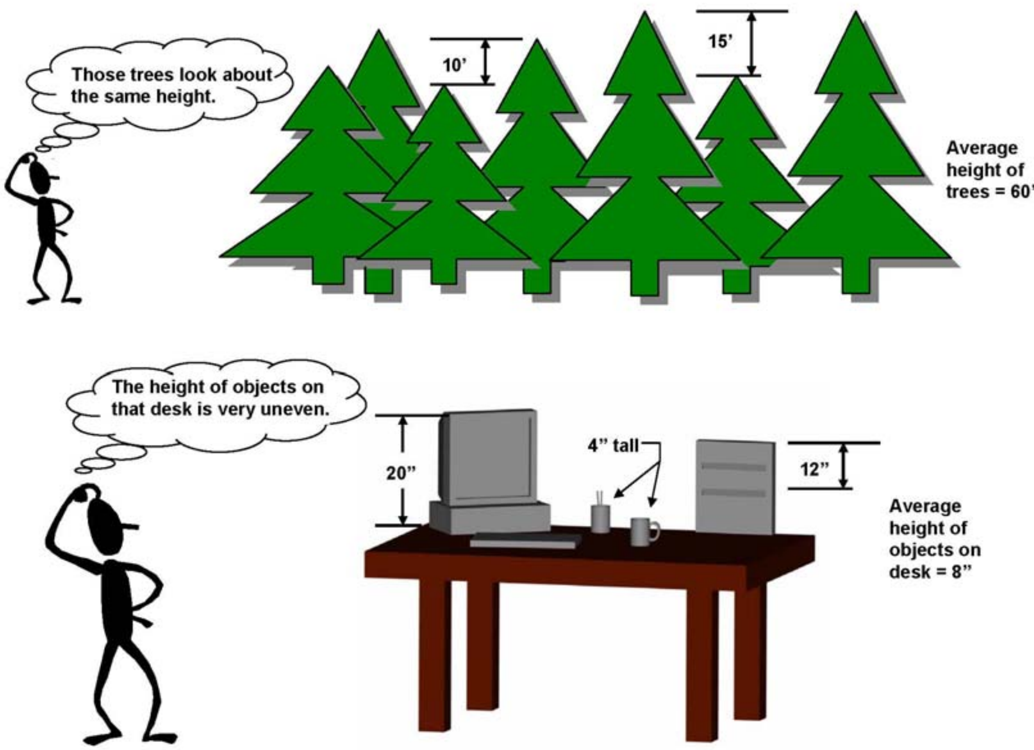

Figure 5.1. Illustration of sensitivity to the coefficient of variation (the standard deviation normalized by the mean) versus absolute measures.

Third, research has also shown that *small probabilities are generally overestimated and large probabilities are underestimated* (Hillel 1973; Hogarth 1975; Sheridan 2002). Figure 5.2 provides a qualitative representation of this from Kahneman and Tversky (1984) and Figure 5.3 provides a logarithmic probability odds scale to help illustrate specific ranges where these overestimations and underestimations have been observed in dramatic fashion. This tendency effects even those who might be initially expected to know better, as Heuer's research revealed: "Intelligence analysts often have difficulty estimating the likelihood of low probability events, especially when those events have potentially very serious consequences" (Heuer 1982, p. 45).

Fourth, as the *number of options is increased, probabilities assigned to similar options changes dramatically*. This was demonstrated by two medical researchers that conducted an experiment to see how physicians respond as the number of possible options for treatment expands. When the doctors seeing a specific patient having a certain symptomology could either be given medication or referred to an orthopedics doctor (i.e., there were only two choices available to the physician: medicine (A) or referral)– the choice was split approximately 50/50. When the scenario was changed such that two medicine options were available in addition to the referral (i.e., three choices were available: medicine (A), medicine (B), or referral), 75% chose the referral (Redelmeier and Shafir 1995; Bernstein 1998).[29] In another example of this effect, 120 Stanford graduates were asked to identify the probabilities associated with dying of natural or unnatural causes. One group of students was presented with a list of explicit forms of natural and unnatural deaths. The second group was given the same task but shown only two category headings: (1) total natural causes and (2) total unnatural causes. People assigned much higher probabilities to natural causes when the multiple, explicit categories were provided (i.e., 73% natural and 53% unnatural for the first group; and 58% natural and 32% unnatural for the second group; actual data at the time showed the true values to be 92% natural and 8% unnatural[30]). The students greatly overestimated the probabilities of violent deaths and underestimated deaths from natural causes (Tversky and Koehler 1994).

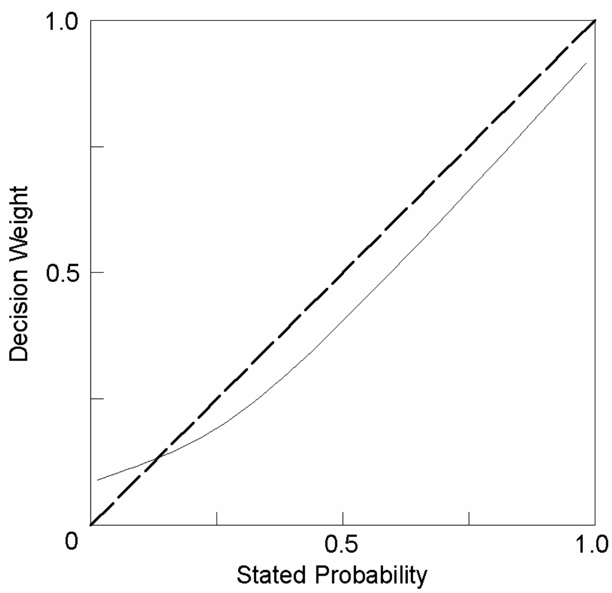

Figure 5.2. Hypothetical weighting function. Taken from Kahneman and Tversky (1984).

[29] In this example, it may be the case that increasing the number of options implies an increased complexity which encourages physicians to seek a second opinion from someone who may have more experience with this complexity.
[30] Note that the probabilities listed by the Stanford graduates do not add up to 100% in either case. This indicates a significant deficit in normative knowledge (i.e., combinatorics, probability theory, statistics and related critical thinking skills).

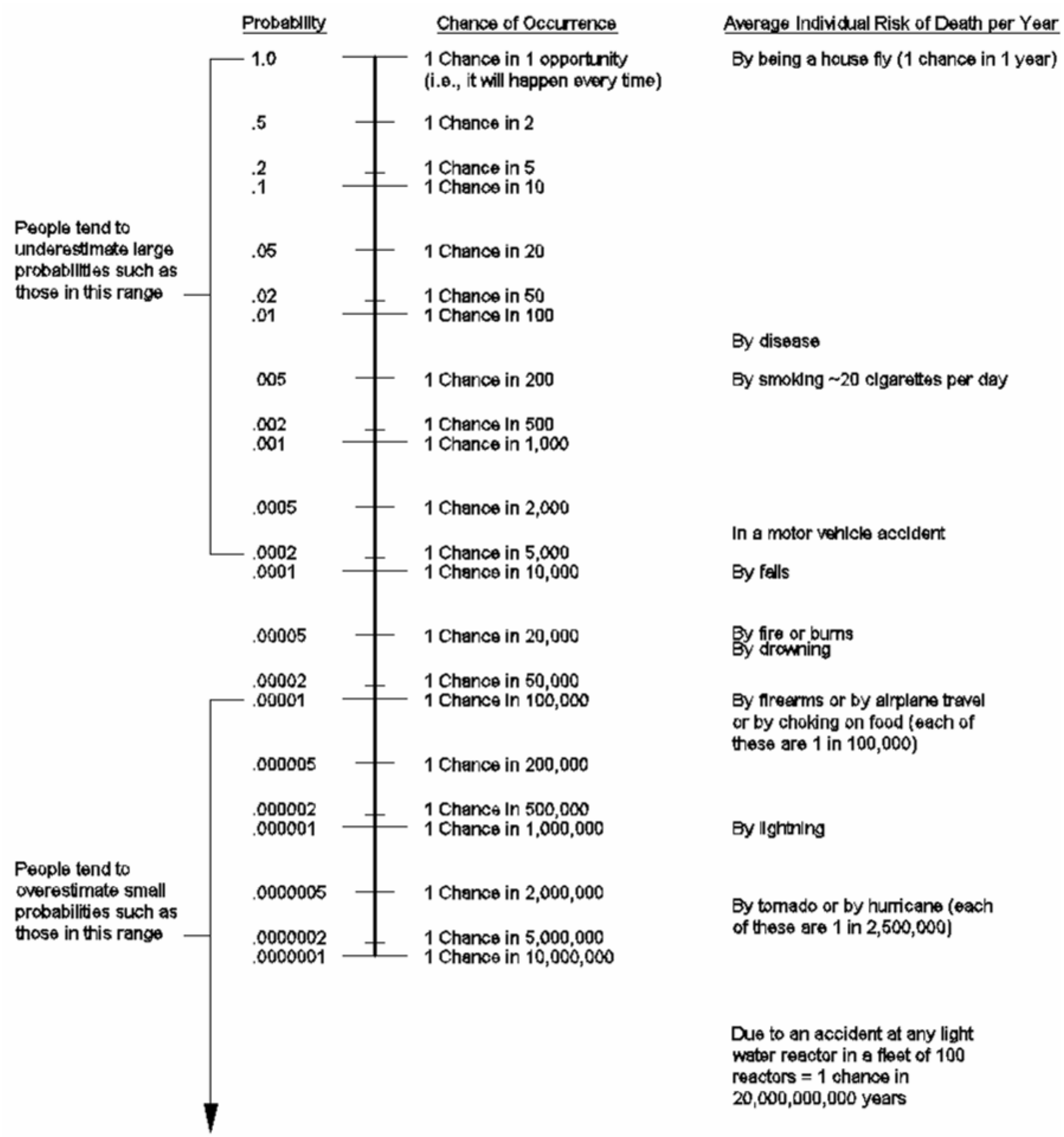

Figure 5.3. Logarithmic probability odds scale with data from Rasmussen (1975) and Saunders and Wade (1983).

A fifth factor that could have played a role in the differences in assigned probabilities between explicit/non-explicit listings of types of death is the level of detail in the descriptions. Tversky and Koehler (1994) state "probability judgments are attached not to events but to descriptions of events…the judged probability of an event depends upon the *explicitness* of its description." This could be seen as analogous to the internal experience of that 'event.' One who is able to imagine the event in detail will tend to attach more weight to that event. Very explicit descriptions in spoken, written, or visual form will enhance the intensity of the experience of that description, i.e., occupy more attention resources, and encourage cognitive 'replaying' of the description multiple times; this strengthens the coding of that explicit description in long-term memory (Kandel, Kupfermann et al. 2000).

Examples of explicit event descriptions that might draw attention for easy future recall include images of mushroom clouds from nuclear bomb blasts (most people have seen, and readily recall this explicit image). Photographic images of the blue glow of nuclear reactor cores would also be very explicit. Hearing the statement "…there is enough radioactive material in one nuclear power reactor core to make many nuclear weapons" might constitute a very explicit message, especially if the listener had memories of nuclear bomb blasts and glowing reactor cores. Another example of an explicit description would be hearing the phrase, "…all traitors of our nation will be shot on site" while simultaneously being shown the very famous Pulitzer-prize winning photograph of a South Vietnamese police officer executing a prisoner with a pistol shot to the head during the Vietnam War. Biasing due to the use of explicit descriptions ties in directly with the availability biases to be introduced later in this section.

A sixth likelihood bias, which is closely tied to the fifth bias, is due to the *representativeness* heuristic (Tversky and Kahneman 1974). That is, people will associate the probability (likelihood) of A belonging to class B, or of A being from process B by the similarity between A & B. An experiment showing how this works includes the following: subjects were presented with a paragraph description of a shy, yet very intelligent person after which they were asked if that person was a farmer, librarian, pilot, salesman, or physician. The subjects ignored the prior probabilities associated with the different occupations, that is they didn't consider that there are more farmers than librarian's, etc. The majority picked librarian as the occupation. When a similar group was asked the same question, even when provided with the proportions of those job categories in the population, the majority still picked librarian as the shy person's occupation. The subjects did use the probabilities correctly when no other information was given. Continuing studies by Tversky and Kahneman showed that even when worthless information was provided, the prior probabilities were ignored. This suggests that people will start digging for representativeness at the expense of probability information–even when they are not competent to judge; all that is needed is for them to *believe* that they are a competent judge. In the same series of studies, it was again shown that people are insensitive to sample size.

The seventh probability estimation bias reported by Fox and Tversky (1995) is called the *ambiguity aversion* bias. They designed a series of experiments to discover whether people's preference for clear over vague probabilities appears in all instances or only in games of chance. "The answer came back loud and clear: people will bet on vague beliefs in situations where they feel especially competent or knowledgeable, but they prefer to bet on chance when they do not" (Bernstein 1998, p. 281). Therefore, the impact of the ambiguity aversion bias is especially sensitive to overconfidence in the decision maker. An example of this in the nuclear power domain could include an anti-nuclear environmentalist who feels that any type of NPP radiation release is qualitatively bad since they read somewhere that there is a non-zero risk associated with even small amounts of radiation. Confronted with a decision, they are likely to use this vague information to support their view instead of accepting clear probabilities showing the low risks of nuclear power versus fossil fuels as determined by systems safety or radiation protection experts. If that same person was confronted with a decision on whether to buy car A or car B based on safety-related probability data, they would likely let the data provide the answer (e.g., evaluations by consumer reports). Of course, in this case the desire for safety may be over-ridden by a desire to drive a 'cute,' energy efficient Volkswagen van. Ambiguity aversion is driven by the feeling of incompetence.

The eighth bias is the well-known *gambler's fallacy*: after observing a long run of black on the roulette wheel, most people erroneously believe that red is now due, presumably because the occurrence of red will result in a more representative sequence than the occurrence of an additional black. Chance is commonly viewed as a self-correcting process in which a deviation in one direction induces a deviation in the opposite direction to restore the equilibrium. In fact, deviations are not "corrected" as a chance process unfolds, they are merely diluted (Bernstein 1998). In other words, long runs of black or red on the roulette wheel can and will happen and the longer one plays, more of these interesting (or frustrating) long runs will be observed. Recall the *weak law of large numbers.* This bias could be applied to the nuclear power industry by someone if they look at the strong safety record of NPPs and decide that after 30 years of safe operation a severe accident is bound to occur soon.[31]

The ninth bias is actually a pair of related biases. People tend to *overestimate the probability of conjunctive events* (due to anchoring on simple individual probabilities) that leads to unwarranted optimism in the evaluation of the likelihood that a plan will succeed or that a project will be completed on time. Individuals also tend to *underestimate the likelihood of disjunctive events*. For example, if the probability of winning a bet is 90% and a person is asked what the probability is for winning the same bet 7 times (i.e., conjunctive set of events–these involve series combinations), they will tend to state a likelihood value that is higher than the actual value of 48% (i.e., $(0.9)^7$). If a person is asked the probability that they will lose at least once with the same type of bet over 7 tries (i.e., disjunctive events–these involve parallel combinations; probability of winning is 90% on each try in this example), they will likely estimate the probability of a loss at less than the actual value of 52%, which is equal to one minus 0.48 (Hillel 1973; Tversky and Kahneman 1974; Sheridan 2002).

The tenth bias observed with probabilistic decisions is that decision makers do not give as much weight to outcomes as suggested by Bayes' rule (Edwards 1968). Probabilities of alternative options tend to be estimated more conservatively (Sheridan 2002). This observation is important given the many attempts by decision making researchers to devise a method for 'optimal' decision making. Bayes' rule is a very helpful approach for aggregating information for use in a decision process, but for those uninitiated in or out of practice in probability theory, its development may be a bit tedious. Therefore, the Appendix to this report presents a summary of several key combinatoric operations along with a simple exposition of Bayes' Theorem showing how a rational information updating system can be derived (Winterfeldt and Edwards 1986; Sheridan 2002). The material in the appendix will be beneficial to the interested/diligent reader, but omitting review of that section will not hamper understandings of any of the other biases.

---

[31] It is important to note that someone who is aware that many of the NPPs in the United States have been operating for close to or over 40 years may be concerned with equipment failures related to aging/wear effects if they do not have confidence in the inspection, test, maintenance, and operational practices used by the industry.

For those readers who do go through the machinations of the Bayesian example in the Appendix, it may be apparent why such an approach is not readily used in the course of human decision making. It requires both specific training and experience to apply and generally requires a disciplined accumulation of data. Given that data do not support the hypothesis that an innate Bayesian estimation tendency exists in people, one must exert considerable effort in mastering the technique. Yet, it is argued here that Bayesian techniques are a very helpful way in which to make maximum use of the data that are available for making a decision and should be the standard against which people's inferential tendencies may be measured (Sheridan and Ferrell 1974). The interested reader is encouraged to consult Winterfeldt and Edwards (1986) for an accessible treatment of Bayesian statistics versus classical statistics[32] and applications of Bayesian techniques to the decision making realm. Another helpful discussion (and easy to follow specific examples) contrasting Bayesian versus classical approaches used in probabilistic risk assessment analyses can be found in NUREG/CR-6823 (2003).

It should come as little surprise that statistical principles are not learned readily from everyday experience since immediate use of information in the environment does not generally require encoding in the form of distributions. For example, people do not discover that successive lines in a text differ more in average word length than do successive pages, because they simply do not attend to the average word length of individual lines or pages. Thus people do not generally learn the relation between sample size and sampling variability, although the data for such learning are abundant (Tversky and Kahneman 1974). The previous discussion of likelihood biases strongly supports this conclusion. It may be speculated that short term memory limitations and a general lack of understanding regarding the power of recording distributions (i.e., a deficit in normative knowledge and experience) encourages a disregard for 'population data' in a person's daily interactions.

Now we address a series of additional biases that also contribute to the ten tendencies mentioned above. Some of these are less closely associated with determining specific probability values, but they are still critical in any decision making process.

One bias deals with the concepts of costs and loss. People tend to treat *costs* and *uncompensated losses* differently, even though their impact is identical. For example, if a person is faced with a gamble involving sums of money they consider to be significant, most people will reject a fair gamble in favor of a certain gain. That is, $100,000 for certain is preferable to a 50-50 chance of getting $200,000 or nothing. Also, when asked whether to pick between an 80% chance of winning $4,000 and a 20% chance of winning nothing versus a 100% chance of receiving $3,000. People chose the $3,000 rather than the $3,200 expectation (i.e., $4,000(.8)). When Kahneman and Tversky reversed the situation by having people choose between taking the risk of an 80% chance of losing $4,000 and a 20% chance of breaking even versus a 100% chance of losing $3,000, now 92% of the respondents chose the gamble" (Kahneman and Tversky 1979).

[32] The author of this paper supports the use of Bayes' Theorem and its extensions as an appropriate method for quantifying *well-structured* inferences. See Winterfeldt and Edwards (1986), Sheridan and Ferrell (1974) and NUREG/CR-6823.

In other words, people with money in their pockets will choose the gamble and people who start out with empty pockets will reject the gamble (Bernstein 1998).

This behavior, although understandable, is not consistent with the assumptions of rational behavior. A person should answer a question in the same way regardless of the setting in which it is posed (Bernstein 1998). This highlights two important items: first, people are not overly risk-averse, they are perfectly willing to choose a gamble when they feel it is appropriate (Kahneman and Tversky 1979; Bernstein 1998). As stated earlier, people are not risk averse, they are *loss averse*.[33] Second, in addition to hating to lose, people are highly swayed by the *framing effect*. This phenomena, regardless of bare facts, means that word choices, image choices, and all aspects of presentation greatly influence the resulting interpretation (Tversky and Kahneman 1981).[34]

To further expand on this concept of loss versus cost there appears to be an interesting mental accounting process used by many people to make such determinations. For example, Kahneman and Tversky (1979) found that a person presented with the situation in which they had purchased a $40 ticket for a Broadway show, traveled to the theater and realized that they had lost the ticket; most people would be reluctant to spend $40 to replace the lost ticket, while about the same number would be perfectly willing to lay out a second $40 to buy the ticket if they had lost the original $40. The important distinction here is that in the first case, the person lost the ticket whereas in the second case they lost money from their wallet. Upon further questioning it appears that in the first case people envisioned the cognitively contrived ‘Broadway’ account as being debited $40 and then $40 again to replace the lost ticket. This resulted in a perception of ‘paying the cost’ of $80 to see the show. In the second case, it appeared that people envisioned the cognitively contrived ‘general money fund’ where they hadn’t really lost anything specific to the Broadway show. Thus the stage show would still be perceived to cost $40 in addition to the separate perception of losing $40. Both events would result in an overall decrease in wealth of $80, but the perception of loss is magnified by drawing on the same mental account relative to a perceived ‘gain,’ which in this case is attending the Broadway show.[35]

Upon closer reflection, the reader may observe that this discussion of costs, losses, and the framing effect related to monetary transactions can be connected to biases presented earlier, such as the bias of focusing on the *coefficient of variation*. For example, people do not tend to think of relatively small monetary transactions in terms of overall wealth, but rather in terms of gains, losses, and neutral outcomes (i.e., maintaining the status quo), that are normalized by the overall amounts of money drawn from a particular ‘mental account.’ Kahneman and Taversky (1984) argue from introspection and

---

[33] That is, people are generally willing to take chances in order to avoid loss.

[34] The *framing effect* is of critical importance when designing, administering and interpreting survey data (e.g., surveys conducted by those supporting the expansion of nuclear power, or surveys conducted by those seeking opposition to nuclear power versus surveys investigating public attitudes toward nuclear power conducted by a disinterested polling group).

[35] It should be noted that this experiment was a hypothetical thought experiment. It is possible that if travel to the Broadway show was difficult or costly, individuals might readily produce another $40 for the ticket. The importance of this example is in how mental accounting pointedly illustrates a subtle aspect of the *loss aversion* phenomenon.

psychophysical measurements that subjective value is a concave function of the size of a monetary gain (see the upper right quadrant of Figure 5.4). The difference between a loss of $300 and a loss of $100 appears greater than the difference in subjective value between a loss of $1,300 and a loss of $1,100 (recall the example of trees in a forest and items on a desktop relative to perceived height changes). Combining this insight with the evidence that people are loss averse leads to a qualitative insight on the shape of a value function (see Figure 5.4).

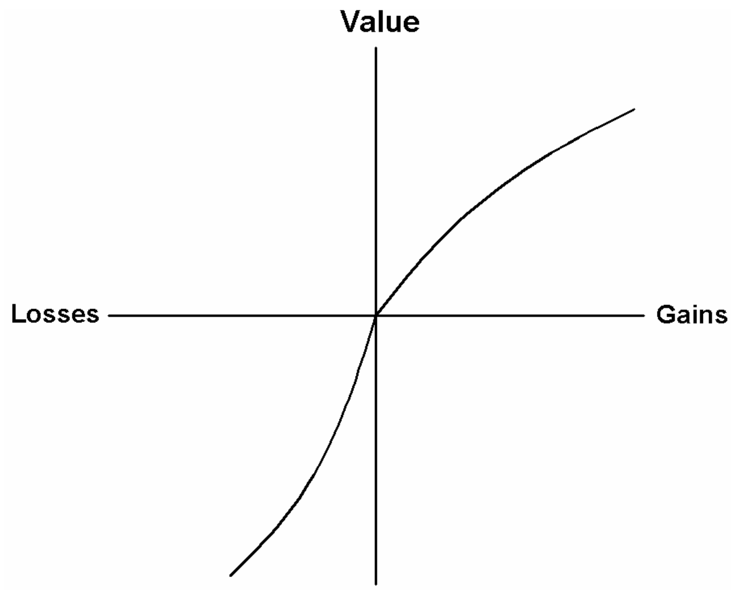

Figure 5.4. Hypothetical value function. Taken from Kahneman and Taversky (1984).

Figure 5.4 illustrates that perception of value is predominantly based on gains and losses,[36] and that the gain of $x is less attractive than a loss of $x is aversive. Stated differently, the beneficial value of the possible gain, is not nearly sufficient to offset the negative value of the possible loss. For example, a large majority of respondents in a sample of undergraduate students refused to stake $10 on the toss of a fair coin if they stood to win less than $30 (Kahneman and Tversky 1984). Is it unusual that people are relatively risk averse in the domain of gains and risk seeking in the domain of losses? That is, people will not take as many gambles to 'gain' as they will to 'avoid loss.' To reinforce this insight, another monetary example (Kahneman and Tversky 1984): consider the case where an individual is forced to choose between an 85% chance to lose $1,000 (with a 15% chance to lose nothing) and a sure loss of $800. A large majority of people will prefer the gamble over the sure loss. This is a risk seeking choice because the expectation of the gamble ((-$1,000)(.85) + ($0)(.15) = -$850) is inferior to the expectation of the sure loss ($800). Risk seeking in the domain of 'perceived losses' has been confirmed by many investigators (Fishburn and Kochenberger 1979; Hershey and Schoemaker 1980; Slovic, Fischhoff et al. 1982). It has also been observed with

[36] Except in the case of possible financial ruin.

nonmonetary outcomes, such as hours of pain (Erakar and Sox 1981) and loss of human lives (Tversky 1977; Tversky and Kahneman 1981; Fischhoff 1983).

The asymmetrical S-shaped function in Figure 5.4 becomes a useful qualitative tool to examine problems with traditional utility determinations in decision theory. The work of von Neumann and Morgenstern (1944) spurred the development of modern decision theory and introduced several axioms that ought to govern the preferences of a decision maker (Kahneman and Tversky 1984). The principal axioms included transitivity (if A is preferred to B and B is preferred to C, then A is to be preferred over C), and substitution (if A is preferred to B, then an even chance to obtain A or C is preferred to an even chance to obtain B or C). Recalling our discussions of the perception of losses and gains, and the framing effect, the reader may notice a problem.

In theory, presenting two versions of a decision problem that are recognized to be equivalent (e.g., in terms of monetary expected value) when shown together should elicit the same preference when shown separately, and minor changes in wording should not matter. Alas, these assumptions for decision making behavior generally do not hold. The following example is taken from Kahneman and Tversky (1984, p. 343) in order to once again reinforce the subtleties of perceived gain and perceived loss (note: items in italics have been added to the original presentation of this example for clarification):

> Problem 1 (N=152, *i.e., sample of 152 people*): Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:
>
> If Program A is adopted, 200 people will be saved. (72% *of respondents picked this option*)
>
> If Program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved. (28% *of respondents picked this option*)
>
> Which of the two programs would you favor?
>
> The formulation of problem 1 implicitly adopts as a reference point a state of affairs in which the disease is allowed to take its toll of 600 lives. The outcomes of the programs include the reference state and two possible gains, measured by the number of lives saved. As expected, preferences are risk averse: A clear majority of respondents prefer saving 200 lives for sure over a gamble that offers a one-third chance of saving 600 lives. Now consider another problem in which the same cover story is followed by a different description of the prospects associated with the two programs:
>
> Problem 2 (N=155, *i.e., sample of 155 people*): If Program C is adopted, 400 people will die. (22% *of respondents picked this option*)

If Program D is adopted, there is a one-third probability that nobody will die and a two-thirds probability that 600 people will die. (78% *of respondents picked this option*)

It is easy to verify that options C and D in Problem 2 are undistinguishable in real terms from options A and B in Problem 1, respectively. The second version, however, assumes a reference state in which no one dies of the disease. The best outcome is the maintenance of this state and the alternatives are losses measured by the number of people that will die of the disease. People who evaluate options in these terms are expected to show a risk seeking preference for the gamble (option D) over the sure loss of 400 lives. Indeed, there is more risk seeking in the second version of the problem than there is risk aversion in the first. The failure of invariance is both pervasive and robust. It is as common among sophisticated respondents[37] as among naïve ones, and it is not eliminated even when the same respondents answer both questions within a few minutes.

The previous example should drive home the fact that framing plays a very large role in influencing peoples decisions. This insight is why business students are admonished to consider all financial problems in terms of total assets rather than in terms of gains and losses (Schlaifer 1959). This is also why lobbyists for credit card companies insisted that price differences between cash and credit purchases be labeled as a cash discount rather than a credit card surcharge (Thaler 1980). Could one imagine the effect on issues of national importance if briefings to the U.S. President were consistently framed in terms of losses? The general strategy for avoiding this type of decision making trap is to deliberately frame a decision problem in more than one way (Fischhoff, Slovic et al. 1980). Specific techniques for doing this will be described later in this section.

Another aspect of loss aversion ties closely into the emotion of fear. Many people fear losing to others, that is, there is an underlying assumption (implicit, or explicit) in many situations that zero-sum constraints apply (Cooper and Robinett III 2005). In other words, in order for me to win, someone else must lose. A simple example is the case of a sports team, where at the conclusion of a game, there will be a winner and a loser. One could also perceive themselves to be the 'winner' on a winning team only if they are allowed to run the touchdowns versus supporting the individuals who do. Another example would be the performance review and raise determination processes used in many companies. When a fixed sum of money is available for giving bonuses and raises, some will 'win' more money than others.

The nature of the reference frame used by a particular individual is of critical importance when discussing loss. A person may perceive themselves a winner after playing a game by the rules and at a very high level of physical/mental effort regardless of the final score; or they may feel like a winner if they receive any raise at all irrespective of amounts of money received by peers; or they may feel like a winner if they accomplish all the professional objectives they set and still have a job at the end of the year. One

[37] Note that 'sophisticated' is not defined. It will be argued later in the paper that with appropriate training and systematic application of the recommended decision making approach, one can mitigate the effects of these biases.

important insight for high-level decision makers is the fact that perceived reference frames and time horizons, which greatly affect the sense of loss aversion, can be greatly influenced by the specific manner in which they are presented; this is a restatement of the *framing effect*. For example, one corporation may reward all employees with significant percentages of end of year profits (during profitable years) and then will lay-off large numbers of employees during non-profitable years. A second corporation may reward employees with much more modest percentages of profits during good years, but build up a reserve which keeps the company from letting employees go during lean years. Is winning defined as high compensation awarded in a given year, or stable employment over 30 years? How these characteristics of the two companies are introduced to new employees and reinforced during employment will guide the interpretations of losses and gains. Explicit attention to loss aversion and the framing effect are important when forming a decision making team and carrying out the decision process.

Applying this type of phenomenon to the nuclear power realm, one can conceive of an example in which a person has made a decision that nuclear power is bad and dangerous even though the decision was made with scant or incorrect information. When the person is confronted with more valid information, they may cling to previously held beliefs since they interpret the 'loss' of that belief as worse than the cost associated with holding onto it, i.e., adhering to an uninformed decision that may prevent them from receiving any benefits from, in this example, nuclear power. This reasoning, termed *confirmation bias*, is strongly supported by studies which show that people tend to seek out evidence that confirms their current position and to disregard evidence that conflicts with their current position (Einhorn and Hogarth 1978). In fact, several studies have specifically shown that preliminary hypotheses based on early, relatively impoverished data interfere with the later interpretations of better, more abundant data (Greenwald, Pratkanis et al. 1986; Reason 1990). The tendency to *disregard conflicting evidence* as a corollary to the confirmation bias is pervasive (Anderson and Jacobson 1965; Slovic 1966).[38]

Klein (1998, p. 68) provides an entertaining example of *confirmation bias*:

> Two physicists associated with the Aspen Center for Physics are climbing in the Maroon Bells Wilderness near Aspen, Colorado. While descending, they lose their bearings and come down on the south side of the mountain instead of the north side, near Aspen. They look below them and see what they identify as Crater Lake, which they would have spotted from the trail leading home. One of them remarks that there is a dock on the lake. Crater Lake does not have a dock. The other physicist replies, "They must have built it since we left this morning." A couple of days later they reach home.

Examples of confirmation bias are continually encountered in daily life. The introspective reader may recall several instances in which they demonstrated this behavior. The married reader will undoubtedly recall such instances (or be reminded of

[38] Another example of confirmation bias in the nuclear power realm might be the tendency to discount evidence showing advancements in safety systems based on earlier experience they may have with less safe systems in that industry.

such occasions by their spouse). The author recalls many such instances; several that come to mind easily happen to involve land navigation on reconnaissance patrols. Others involve urban roadway navigation incidents, which the author's wife remembers even better than the author does (e.g., I know where that short cut is, it's right here…).

Another important example of confirmation bias helps to tie this phenomenon in with the previously mentioned critical thinking process number eight, which called for discrimination between inductive and deductive reasoning. Research regarding the intelligence analysis task of deciphering blurry images has shown that the more blurry an initial image, the longer it will take for identification. That is, the identification time is a function of blurriness such that the greater the initial blur, the clearer the image had to be before subjects could recognize the image. When blurring is great, people have more time to formulate initial hypotheses that must be eliminated as the correct image becomes apparent. Subjects not exposed to the blurry image recognized the true image much faster than those who struggled first with the blurry versions. Intelligence analysts routinely examine new problems at early stages where information is fuzzy and must use structured, disciplined techniques to combat biases (Heuer 1982).

The next few biases are very closely related under the heading of the ***availability heuristic***. They each involve the way in which experiences or thoughts are instantiated,[39] stored into memory and then retrieved with respect to a decision making event. These biases include *retrievability, salience, imaginability,* and *illusory correlation* (Tversky and Kahneman 1974). The concept of *availability* as a judgmental heuristic refers to the ease with which things can be brought to mind. For example, one may asses the risk of heart attack among middle-aged people by recalling such occurrences among one's acquaintances. Another example would be the assessment of nuclear power plant meltdown risk as being analogous to the ease with which one can remember the last media report of a radiation release or 'risk of meltdown' being reported.

Retrievability, as a type of availability bias involves how easily the item can be brought out of memory, or constructed out of memory-type mental processes. An example of this effect included an experiment in which a list of male and female names were announced to subjects. Each list of names included an equal number of identifiably male or female names, but different groups of subjects were presented with lists containing either a high number of famous males or famous females (e.g., George Washington, Marilyn Monroe, etc.). The groups tended to greatly overestimate the number of names announced that were male or female if the famous names were male or female, respectively (Tversky and Kahneman 1974). The reader should note the important point that 'retrieving' something complex from memory does not actually get represented verbatim each time. Parts of the complex memory will be reconstructed from key events and noticeable variations will occur each time. Also, the frequency with which complex items are consciously recalled will have a strong effect on how persistent and consistent memories of the key events will be, i.e., while long-term memory has no known capacity limitations, particular memories do decay over time (Kandel, Kupfermann et al. 2000).

[39] Instantiated is defined in Webster's as: to represent (an abstraction) by a concrete instance (Mish 1996). In this case, the word is further being used to explain the general process of perception and assignment of meaning that occurs during intentionally directed person-environment interactions.

Another example of retrievability would be the situation in which an airline pilot sees a low fuel quantity indication during flight, having experienced several previous instances of faulty fuel quantity indicators, the hypothesis of a faulty gauge receives excessive attention during his troubleshooting activities. The chance that fueling personnel made a fuel loading error, or the existence of a major fuel leak and failure to notice rapidly increasing automated trim corrections are being made by the flight management system to keep the right wing level[40] are discounted until one of the engines finally shuts off due to fuel loss. Recall the critical thinking processes 1, 3, 5, 6 and 7. Consider also the time constraints that may prevent an airline pilot from systematically applying those processes before being overwhelmed by system events. Temporal constraints related to identification and mitigation of biases using critical thinking and the well-structured decision making strategies can be a major stumbling block in cockpit traveling ~600 miles per hour. Fortunately, such extreme temporal constraints are not present in high consequence, highly technological decisions such as how to structure major elements of the energy infrastructure in the United States to provide reliable, safe, and sustainable sources of electricity and transportation fuels.

When discussing retrievability, and for that matter all of the other availability-type biases, it is essential to mention limitations on working memory. There is a finite capacity of the consciously controlled memory space in which disparate ideas or other pieces/chunks of information can be manipulated. Miller (1956) conducted an extensive review of experiments involving memory processes and he found that in general, people were capable of consciously storing and manipulating 5 to 9 pieces of information in what is now referred to as working memory. Anyone who has studied memory processes will have encountered references to Miller's (1956) paper and will probably recall its very memorable title: *The magical number seven plus or minus two: Some limits on our capacity for processing information* (or at least the first 8 words of the title will be recalled). Fairly complex 'chunks' of information can be included within those 5–9 memory items if structured techniques are followed. See Wickens and Hollands (2000) for a good discussion.

In the domain of high-consequence, technologically advanced systems decision making often involves generating mental simulations or logical arguments for which aspects of the system need to be considered, speculations on how 'important' parts of the system may interact, etc. How exactly does the 5–9 item working memory limit impact this process? To answer that question we turn to research by Klein and Crandall (1995) on naturalistic decision making.[41] They screened many documented examples of decision making in critical situations and arrived at seventy-nine cases that clearly revealed evidence of mental simulation. That is, the decision maker consciously simulated their way through the problem at hand. For example, the thought processes may proceed as follows, "Here's the starting point. Then this happens, which leads to this event, then this other thing kicks in, and you end up right here."

---

[40] Corrections to keep the right wing level would be made to compensate for the reduced mass of the wing as fuel is expended.

[41] In this case, naturalistic decision making involves an experienced decision maker (e.g., fire fighter, fire team commander, military personnel, etc.) who is 'thinking up' one or more mental simulations in order to solve a problem that needs to be solved in a relatively short period of time.

Klein and Crandall were intrigued to discover that the mental simulations they studied were not very elaborate. Each was relying on just a few factors–rarely did they see more than three. This would be analogous to mentally designing a machine with only three moving parts. Another striking feature was found; the mental simulations rarely moved through more than six transition states of the 'imagined' system. It was speculated that working memory limitations may be the reason for mental simulations being completed in roughly six steps (Klein and Crandall 1995). This has direct application to the high-consequence, technologically advanced decision making in the following way: during the decision making process, if the decision making team finds that the set of decision pathways/options analyses move through six or less transition steps, it is highly likely that human working memory limitations are unnecessarily restricting the analysis. For example, when considering "if this, …then this" aspects of the decision domain, be sure to use additional means of memory storage (e.g., writing things down, computer-based tools, etc.) to make better use of analytical capabilities to handle system complexity.

At this point, the clever reader may ask: But what about the principle of parsimony? Aren't we struggling for the simplest understanding/theory/model of the decision domain that we can find? Won't that generally keep us in the region of models we can manipulate in working memory? The simple answer to this question in the realm of high-consequence, technologically advanced decision domains is **no**. Here's why. Occam's Razor (also spelled Ockham's Razor), attributed to the 14th-century English logician and Franciscan friar, William of Ockham, has unfortunately been reduced in many people's minds to the statement: the simplest explanation is the best. This basic definition is problematic as it leads many to the invalid heuristic which deems theories/models to be better merely by virtue of being easier to comprehend. The proper understanding of the principle of parsimony (a.k.a., the law of economy) is to interpret simplicity as the lack of unnecessarily improbable loose ends in a theory or model in favor of a theory or model that does a better "job" of tying things up. Stated differently: when deciding between two models that make *equivalent* predictions, choose the simpler one (Sober 1981; Sober 1990; Nolan 1997; Wikipedia 2005).

Continuing on, when considering high-consequence, technologically advanced decisions it is important to recall the Pareto Principle. In nineteenth-century Italy, the Italian economist Vilfredo Pareto observed that roughly 80 percent of the country's wealth was controlled by roughly 20 percent of the population. This insight lead to what is called the Pareto Principle or the "80–20" rule (Kolarik 1995). This concept has subsequently been applied to many areas, such as industrial system improvements (Juran and Gryna 1988; Juran 1989). For high-consequence systems the Pareto Principle may be interpreted to suggest that the majority of undesirable system outcomes are distributed such that a 'vital few' problems always constitute a high percentage of the overall negative outcomes (not necessarily in an 80–20 manner, it could be 70–30, 90–10, etc.). Unfortunately for the decision maker, high-consequence systems have two critical features: first, there is an extremely low tolerance for undesirable outcomes (recall nuclear weapon detonations, major reactor accidents, major electricity and transportation energy shortfalls, etc.); second, there is very little observational data (or no data) with which to probe the existing or envisioned system (i.e., many known unknowns, and unknown unknowns). Therefore, the analysis of the decision domain must aggressively seek understandings for many of the factors hiding in the 10–30 percent 'tail' of a Pareto tabulation of factors resulting in undesirable outcomes. The lesson: for high consequence systems, beware of oversimplified models and apply Occam's Razor as it was intended.

Another aspect of availability is referred to as *salience*. This is associated with the stimulation of the senses and how strong sensory input demands attention resources. For example, having seen a house burn up in flames with the associated dramatic, attention-demanding explosions and the loud crackling of furnishings being consumed as fuel will have more of an impact on a person's assessment of the likelihood of house fires than if they had read an event description in a newspaper. Presumably, this is due to the high level of attention that the sensory stimuli demanded based on perceived dangers and/or consequence. It is likely that both innate and learned behaviors contribute to salience levels.[42] Wickens and Hollands (2000) provide extensive discussion on attention resources in their exposition of the information processing description of human cognition and human-environment interaction. Figure 5.5 shows an example of a collage that many would agree as having visual salience–irrespective of whether one tends to support or oppose nuclear power–it commands attention.

Figure 5.5. An example of a salient graphic, which implies that media organizations often depict radiation as a frightening phenomenon.

In addition, the availability heuristic also includes *imaginability*. Tversky and Kahneman (1974) found that people tend to generate several instances of events from memory and

[42] Another example of a salient stimulus would be an audible fire alarm that is tuned to a resonant frequency in the human ear. The noise and discomfort of the warning will readily command conscious attention and motivate many people to take some type of action.

evaluate the frequency or probability of occurrence based on the ease with which these events can be constructed. As alluded to earlier in this report, the term ‘constructed’ is very appropriate as retrieval of memories involves bringing different kinds of information together from different locations within the central nervous system. This mechanism is nearly identical in form to the processes that are involved in perception of external events. Therefore, the constructive memory process is subject to distortion just as sensory perception is subject to illusions. Furthermore, retrieval of information is most effective if it occurs in the same context and with the same cues that were present when the event was learned or otherwise experienced[43] (Kandel, Kupfermann et al. 2000). As an example of imaginability, you might have seen ‘hyped’ up media stories about the pervasiveness of ‘mysterious radiation’ or old footage of nuclear bomb tests during the reporting of a minor radiation accident (i.e., a < 1 mrem exposure to the environment), it may then be very easy for you to imagine a ‘release’ of ‘deadly’ radiation or series of releases. Therefore, you may internally assign a much greater likelihood for this event than what is warranted by objective data. This effect is also heightened by *recency*, as recent events are generally easier to recall.

One other specific bias related to the availability heuristic is *illusory correlation*, which involves associating two things together without reflecting on how weak that connection is or should be (Tverskey and Kahneman 1973; Hogarth 1975). This appears when multiple items that are easy to recall at the same time (e.g., they may have been encoded into memory at nearly the same point in time) may be perceived as having a causal relationship. For example, if someone sees two people with similar facial features and/or skin color introduced during a televised news report as being terrorists caught plotting to steal special nuclear material, that person may then think that the next similar looking person they see is a terrorist. This illusory correlation is very persistent, even when contrary information is presented (Tverskey and Kahneman 1974).

Another interesting bias involves the *anchoring effect.* This bias describes the observation that different starting points for values lead to different estimates, which are biased toward the initial estimate (note the probable connection here with the family of availability biases). For example, when Tversky and Kahneman (1974) showed different groups different starting points, these arbitrary numbers had a marked affect on estimates. Specifically, the median estimates of the percentage of African countries in the United Nations were 25 and 45 for groups that received 10 and 65, respectively, as starting points. Payoffs for accuracy did not reduce the anchoring effect. Anchoring occurs not only when the starting point is given, but also when the subject bases his estimate on the result of some incomplete computation. For instance, 8 x 7 x 6 x 5 x 4 x 3 x 2 x 1 versus 1 x 2 x 3 x 4 x 5 x 6 x 7 x 8 produces different estimates of what the final multiplication value will be when subjects have only 5 seconds to register their guess. Subjects estimated 2,250 (median value) as the product of the first sequence and 512 (median value) as the product of the second sequence, while the actual answer is

[43] To highlight how important specific cues are, recall a visit to a once familiar place; e.g., the house in which you grew up. Visiting that place later in life will bring back a flood of memories that correspond more correctly to the actual events that took place many years before than if the reflection occurs elsewhere.

40,320 (Tversky and Kahneman 1974). To summarize the anchoring effect, people are biased toward the first option they see or the first judgment they make.[44]

The final bias to be discussed in this section is the very famous *hindsight bias*, in which a person recalls having greater confidence in an outcome's occurrence or lack of occurrence than they had before the resulting events were known (Fischoff 1975; Sheridan 2002). An example of this bias would include the following scenario: a person who identified himself as being anti-nuclear due to fears of radiation expressed opposition to nuclear power plants during the period from 1972–1977. This person was concerned about the risks of nuclear power, but was not overly frightened due to the fact that many years of operation had passed without a major incident. After watching the exaggerated reports of impending doom during the TMI accident in 1979 this person rapidly experiences a distortion of memory such that in a conversation with a close friend he insisted that he 'knew' a major accident had been imminent for years.

The previous discussion of confirmation bias, framing effects, loss aversion, availability biases, and the myriad of other biases, may have painted a relatively bleak picture for the chance of successfully mitigating these powerful biases using a structured decision making approach. In fact, some might argue that you cannot teach people how to overcome such biases. Fortunately, there is evidence to the contrary. The position taken in this report is that one certainly can use a structured process to learn foundational critical thinking processes and to mitigate the above-mentioned biases in order to make the best use of available data. In developing a line of support for this position, let us first borrow insights from visual perception. Figure 5.6 provides a selection of very famous biases or tendencies for interpretation of visual information.

In Figure 5.6 (A) we see figure-ground recognition, sometimes you see a white vase and other times you see a pair of faces–similar to the signal versus noise distinction in communication theory; the perception of seeing both images is actually due to a rapid shifting from one interpretation to the other. Figure 5.6 (B) shows that one's perception of size is heavily dependent upon the other objects in view, the image of the smaller woman sitting is identical in size in both left and right panels (measure them if you are not convinced). Figure 5.6 (C) demonstrates powerful assumptions regarding the origin of ambient lighting. Most people will initially assume that spheres are being lighted from above, but if one assumes that light is coming from below, then spherical cavities are 'seen.' In Figure 5.6 (D) try to imagine that both columns are either convex or concave–it's nearly impossible to do so. Figure 5.6 (E) displays the Kanizsa triangle, which emerges due to the assumption of continuous shapes–an appropriate assumption in many real situations. Figure 5.6 (F) appears to be a random collection of shapes scattered about, and Figure 5.6 (G) adds an ink blob to highlight the brain's assumption of object layering–can you see the multiple letter B's? Finally, Figure 5.6 (H) is the well-known Müller-Lyer illusion. Both of the lines in (H) are identical in length, but the line with arrow heads appears much shorter–due to size versus shape assumptions.

[44] A common example of the anchoring bias would be comparing the success of ones negotiating skill by the amount a car is purchased below the 'sticker' price or the amount a house is purchase below the initial listing price. The anchoring bias is a strong driver behind the success of sales advertising 50% savings or more on inflated 'suggested retail prices.'

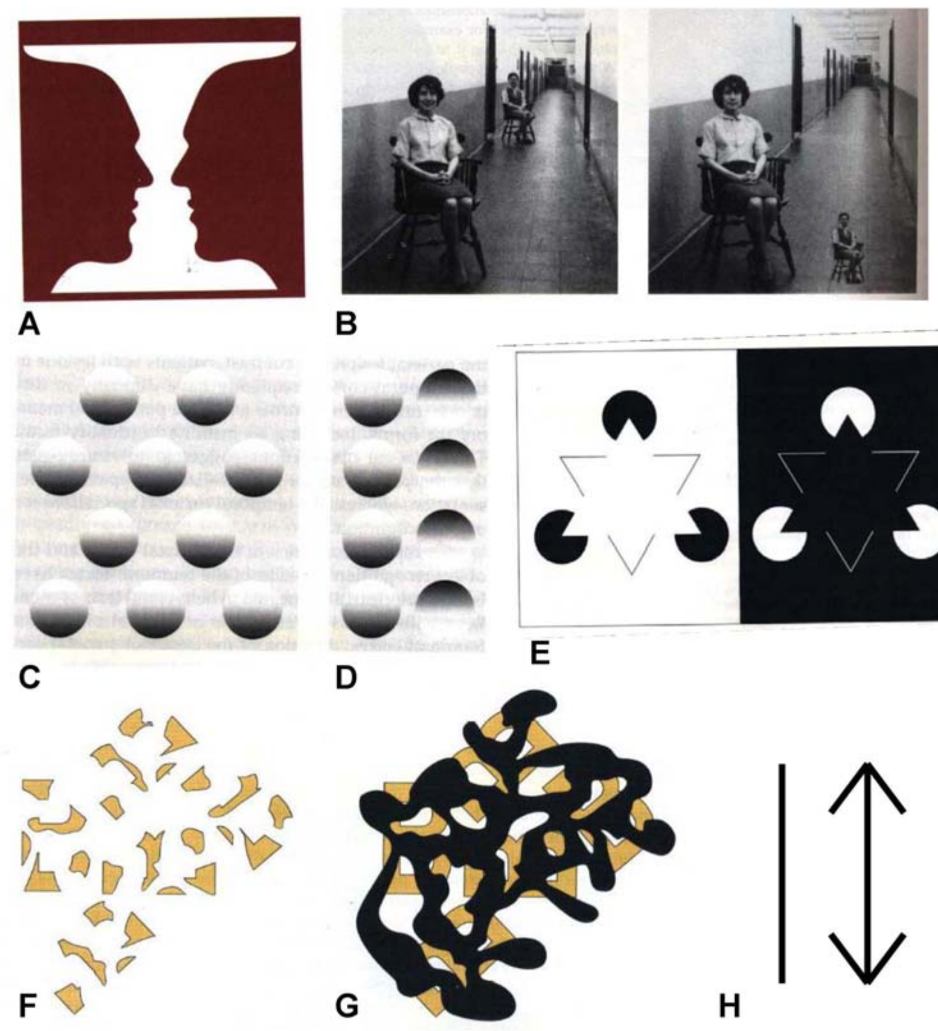

Figure 5.6. Visual illusions. In (A) figure-ground recognition; (B) relative nature of perceived size; (C) and (D) lighting assumptions; (E) Kanizsa triangle; (F) no discernable pattern; (G) assumption of layering reveals multiple letter B's; and (H) the classic Müller-Lyer illusion adapted from several figures in Kandel and Wurtz (2000).

The visual 'biases' in Figure 5.6 are quite compelling, even to those trained in such perceptual phenomena. But is it possible to identify such potential visual illusions and then correctly interpret them? Yes. With proper training, experience and vigilance, one can be wary of these potential illusions when interpreting data and extract invariant pieces of information. How does this relate to the domain of critical thinking processes and decision making biases? *Will* every person who is given proper training and opportunities for first-hand experience master these techniques? No. Some people will not internalize the techniques due to variations in motivation, practice, or vigilance. *Can* every person who is given proper training and opportunities for experience master these techniques? Most certainly can, but a significant percentage likely will not. In the domain of teaching pre- and in-service elementary school teachers and other nonscience majors[45] key concepts in mathematics and physical sciences,[46] Arons (1990) has discovered that the following history of progress is typical: the percentage of successful performers will increase substantially with the first few repetitions of the material, but the curve becomes concave downward and invariably levels off somewhere between 70–90%. Even with intense personal tutoring, roughly 15% of the students never developed the capacity to perform successfully on a particular task. Arons (1990) continues:

> There are certain obvious questions: Were we insufficiently skillful in providing guidance and instruction to the unsuccessful 15%? Would success still be achieved over much longer periods of time? What would have happened if these individuals had received such instruction at the age of 11 or 12 instead of so much later? Are there some individuals who are intrinsically unable to develop these capacities? Are our observations, for some reason, invalid? All we can say is that we do not know the answers. We had reached a point of diminishing returns under available time and resources and were unable to press the issue further. The empirical fact is that the progress curve leveled off below 100%. We hope answers to some of the preceding questions will begin to emerge as time goes by (p. 322).

When similar concepts are presented to science majors, the learning is often faster and a higher percentage will successfully perform the task, but the type of instruction (i.e., inquiry oriented) and degree of repetition is absolutely essential for maintaining the ability over time in a high percentage of the students (Arons 1990). Therefore, it appears that it may be possible to achieve practically significant levels of success in teaching both critical thinking skills and decision making bias mitigation techniques, and providing a practical approach for applying those skills/techniques to a specific decision domain. This report implies ways to teach those general techniques and explicitly provides a method for structuring one's approach to a specific decision domain. The author has developed a process to explicitly teach both the general skills/techniques and the practical approach in addition to measuring decision making performance before and after the instruction.[47]

To support the particular claim that structured approaches can help mitigate decision making biases, here is a technique used by Klein (1998) and his associates to counter the framing effect and confirmation biases. It is called the premortem strategy. The goal is to give decision makers a perspective from which they can aggressively search for flaws in a plan of action they have created. A facilitator asks the decision makers to imagine themselves months or years into the future with the knowledge that their plan has been carried out and that their plan has failed. That is all the information they are given. It is their task to look for causes that lead to statements such as, "Of course, the plan was likely to fail because…" This technique has proven successful in a number of military planning exercises by breaking the strong emotional attachment people have to the plan

[45] Ranging in age from 18 to over 30 years.

[46] For example, a specific task such as arithmetical reasoning involving division; forming a clear operational and intuitive distinction between mass and volume; etc.

[47] The process involves computer-based training modules and a module facilitating application of the structured approach to a specific decision domain. Funding support for the project is being sought.

they have created in a way that lets them demonstrate how creative and competent they are in diagnosing failure.

In an effort to make sense of the many decision making biases/tendencies presented, it is helpful to group the above-mentioned biases into categories and to consider the areas where overlap likely occurs. Figure 5.7 groups the many biases presented earlier into three categories: normative knowledge–characterized by an understanding of combinatorics, probability theory, and statistics; availability–a host of biases that result from how human perception, learning, memory and communication mechanisms function; and individual specific–a person's values, personality traits, interests, associations, and decision domain knowledge.

Figure 5.7 represents the author's attempt at arranging risk perception and decision making tendencies in categories that appear somewhat orthogonal (i.e., normative knowledge, availability biases, and individual specific) in two major respects. First, they are proposed to be different with regard to how easily one may improve their capacities for decision making in those areas. Second, the first two categories (i.e., normative knowledge and availability) comprise a large number of biases that are relatively easy to 'depersonalize' from a specific individual. That is, different people have varying capacities for carrying out normative tasks because of specific training and experience histories. Availability biases result from the fact that we are human beings. Only the individual specific biases are intimately tied to each person's identity and personality.

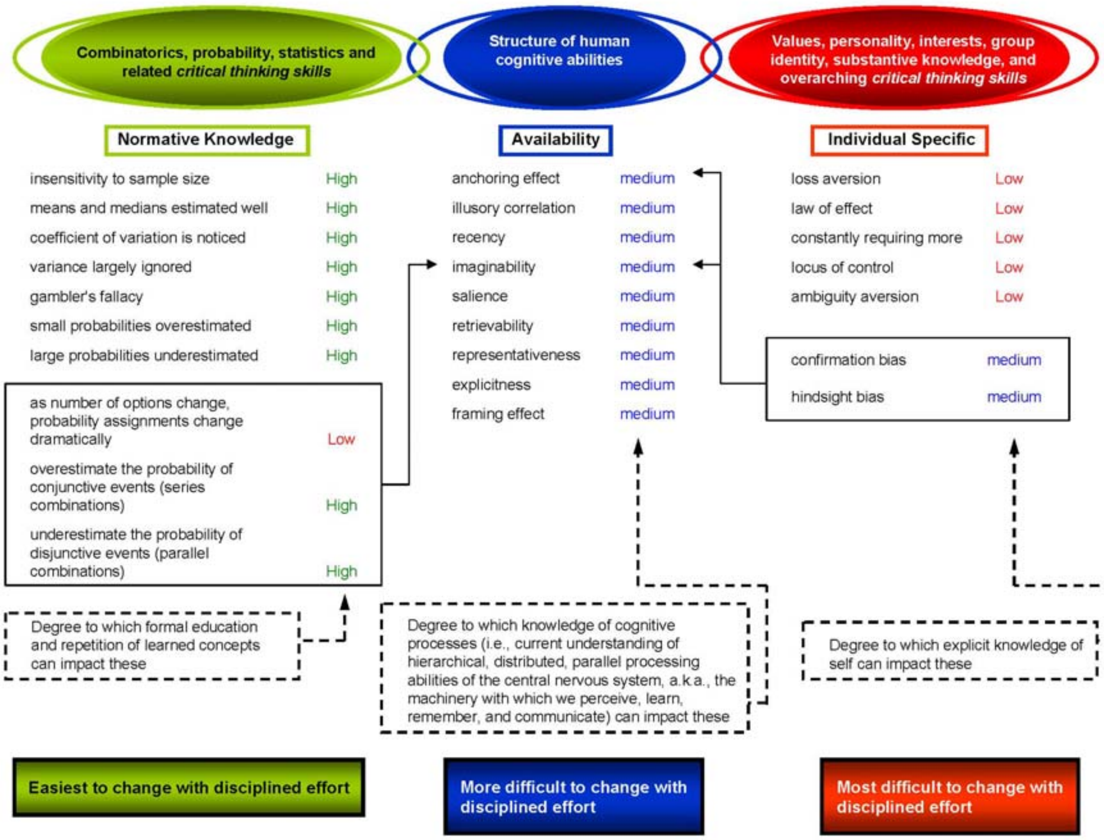

Figure 5.7. The unique framework in which to understand decision making biases/tendencies that researchers have identified. The unique aspects *introduced by the author*, following a review of relevant data, include: 1) the three primary categories into which previously investigated biases are ordered, 2) strong interdependencies hypothesized between biases as indicated by solid lines with arrows, and 3) degrees to which these biases may be mitigated via focused efforts by a specific decision maker indicated by dashed lines with arrows; high, medium, and low ratings;[48] and the green, blue, and red boxed items. This report is intended to present strong arguments supporting the face validity of the unique framework above and the associated decision-making improvement strategies.

Of course, there is considerable overlap among these categories and it should be recognized that the individual specific biases arise in large part from interactions with the environment that are mediated through availability, and normative knowledge filters. Even so, it is sincerely hoped that the unique framework summarized in Figure 5.7 proves to be a useful tool in improving decision making, especially when the critical thinking processes mentioned in section 3 of this report are incorporated into the process. The next section takes all that has been presented and formulates it into a step-by-step decision making approach.

[48] The high, medium, and low ratings should be interpreted as comparative measurements on an *ordinal scale*, i.e., the arrangement represents high-to-low assessments without specific delineation of the intervals separating the assessments. Further research is required to locate the biases on an *interval scale* or a *ratio scale*. See Stevens (1946) and Conover (1999) for a thorough discussion of scales of measurement.

6. A Decision-Making Approach

It is not the purpose of this report to exhaustively cover the complex topic of high-consequence decision making, but it is hoped that the decision making approach presented here will fruitfully bring together the topics discussed thus far into a method that provides the reader with a good means to generate, evaluate, and select among decision options involving high-consequence, advanced technologies. As mentioned earlier, the approach proposed in this report is designed for decision making settings in which a significant amount of time is available to the decision makers. Of course, it is realized that the high-level decision makers who have the authority and responsibility for such high-consequence decisions are often very busy addressing a multitude of important decisions at any one time. Therefore, elements of this approach will likely be applied by trusted subordinates of the responsible decision makers. Ideally, the trusted subordinates will exert tremendous effort to articulate the decision domain and potential decision options in as truthful a manner as possible. Here, truthful manner refers to one in which mitigation strategies are used to avoid unhelpful biases (i.e., those that systematically hinder proper interpretation of available data), and critical thinking processes (as defined in this report) are diligently applied.

To aid in this process, a step-by-step decision making approach is presented in this section. Unfortunately, there is no way to make such a decision process ‘easy’ to apply to a specific domain. Successful critical thinking takes considerable cognitive effort. This report can only provide guidance that helps decision makers navigate away from decision-making traps which ensnare the unwary and provide some grounds for arguing procedural completeness. It remains the responsibility of the decision makers to put forth the effort and exercise due diligence in executing the approach in order that those affected by the resulting decision will derive maximum benefit. Fulfilling such a responsibility requires humility and courage, as faithful application will mean refraining from using knowledge of decision making biases to achieve potentially great personal gain at the expense of others. In simple terms, one must struggle to do the ‘right thing.’

It is expected that a team of decision makers (and trusted subordinates) will be attacking any high-consequence decision, so a team-based example is provided below to emphasize the importance of exerting considerable effort and doing the ‘right thing’ when the stakes are very high:

The example involves the intense practice of immediate action (IA) drills by special operations military teams (e.g., USMC Reconnaissance, USN SEALS, US Army Special Forces, USAF Combat Controllers, etc.).[49] These teams routinely operate in isolated 4–8 man teams far behind enemy lines. IAs are highly disciplined routines for fire and maneuver actions in response to an unexpected enemy attack. Each member of the team has a specific task to execute based upon the direction of the attack, the type of weapons used by the attackers, and any coordinating commands issued (usually non-verbally) by fellow team members. Failure to perform the appropriate IA due to fear, hesitation, or confusion virtually guarantees death for one or more members of the team. These teams are highly disciplined, confident, very well practiced, and each team member trusts each

[49] The author is a former Reconnaissance Marine, hence the biased ordering of military teams in this list.

other team member with their life, in fact, these teams will have proven that trust on multiple occasions merely during training exercises.[50] When attacked, these teams react appropriately.

Similar tactics are taught to regular infantry units. When faced with an unexpected attack, the response of these teams exhibits much more variability. Some teams may have one or more deficiencies in discipline, confidence, training, or trust, which results in a mix of flight, fight, or freeze behavior, or they may actually execute the IA successfully. One critical factor determining success or failure is the degree of previous combat experience. Seasoned small-unit warriors have learned that survival requires teamwork and absolute trust; whether you have known your teammates for three years or three minutes–you must do the 'right thing.' Any team can suffer tremendously from the mistakes of just one inattentive, undisciplined, or generally self-centered team member. The significance of this effect is passed down to new recruits using powerful stories about the infamous 'ten-percenter.' The ten-percenter is the one who fails to do what he is supposed to do when it must be done (i.e., they display 10% performance versus 100%). As the stories are told, it is the ten-percenter who often survives an engagement while teammates doing the 'right thing' are killed. Woe to the soldier, airman, seaman, or Marine who is identified as a 'ten-percenter' in training or combat.

Hopefully, the team-based example above has encouraged the reader to assume a disposition of decision making excellence; we now consider the decision making approach. First, it is important for each member of a high-consequence, decision making team to have some self knowledge regarding items that will have a *general impact* on their decision making performance. Each decision maker should retain control over which items regarding their 'personal' information are shared with other decision makers. It is anticipated that different decision teams will vary widely in the degree to which they voluntarily offer 'personal' perspectives to others. Ideally, strong bonds of trust would form among decision makers during the decision process and sharing of these personal items would be increased. Second, the approach suggests actions to take when considering a *specific* decision domain.

<u>Step 1</u>

Before approaching *any* high-consequence, multi-attribute decision, the responsible decision makers should do the following:

1. Explore what you value most and what you value with respect to other major groups of people by answering these questions:
    A. What is good? (i.e., good for me; good for other groups)
    B. What is bad? (i.e., bad for me; bad for other groups)
    C. What are my personality traits? (esp., habits that relate to decision making)
    D. What group identifications do I have? (i.e., group associations both professional and personal that may strongly shape decisions)

[50] That is, the training exercises involve numerous life-threatening hazards which provide trust-testing opportunities; obviously, combat seasoned teams will have proven that trust many times over in the most treacherous of environments.

2. Honestly reflect on your level of understanding of cognitive capabilities (i.e., heuristics used by humans to simplify decision making in complex environments) by answering the following questions:
    A. What is my training in human perception, learning, memory, and conscious capacities?
    B. Do I need to review my previous training or learn more about these topics before making the decision?
3. Honestly reflect on your level of normative knowledge (i.e., combinatorics, probability, and statistics competency):
    A. What is my training in combinatorics, probability theory, and statistics?
    B. Do I need to review my previous training or learn more about these topics before I make a high-consequence, multi-attribute decision?

<u>Step 2</u>

Make an initial attempt at articulating the decision domain, gathering what appears to be relevant data, and information in a coherent (hopefully written) form, and then systematically screen for places where faults in the ten critical thinking skills and/or the decision making biases/tendencies identified in this report may be playing a role in this initial formulation. Try to frame aspects of the decision domain in multiple ways in order to probe for uncertainties and drive efforts to seek additional clarifying data/information and/or decision team members. This initial attempt to define and understand the decision domain will probably include only a subset of the decision team members who should ultimately contribute to the decision.

When approaching a *specific* high-consequence, multi-attribute decision, the responsible decision maker should do the following:

1. Explore what you value most and what they value relative to other major groups of people with respect to the specific decision at hand by answering these questions:
    A. What is good with respect to potential outcomes of this decision? (i.e., good for you; good for other groups)
    B. What is bad with respect to potential outcomes of this decision? (i.e., bad for you; bad for other groups)
    C. Which of my personality traits are especially relevant to this decision?
        i. What are my general habits that relate to this decision?
        ii. What things do I like about this decision topic?
        iii. What things do I not like about this decision topic?
    D. Group identification? (i.e., group associations both professional and personal that may strongly shape my approach in this decision)
2. Honestly reflect on your level of knowledge or ignorance with the specific decision topic by answering the following questions:
    A. What has formed my understanding of the topic(s) closely related to this decision? (e.g., formal education, discussions with friends & colleagues, scholarly research, personal experiences)
    B. Do I need to review my training or understanding of the topic area?
3. Honestly reflect on your level of understanding of cognitive capabilities (i.e., heuristics used by humans to simplify decision making in complex environments) by answering the following questions:

    A. What is my training in human perception, learning, memory, and conscious capacities?
    B. Do I need to review my previous training or learn more about these topics before making the decision?
4. Honestly reflect on their level of normative knowledge (i.e., combinatorics, probability, and statistics competency):
    A. What is my training in combinatorics, probability theory, and statistics?
    B. Do I need to review my previous training or learn more about these topics before I make a high-consequence, multi-attribute decision?
5. Build a value tree and inference tree specific to the decision at hand.[51]
6. Repeat/review items 1–6 in this step at least once.

Step 3

After one or more iterations of step number 2, reflect on the make up of the decision making team, and add to the team as indicated by deficiencies in normative knowledge, deficiencies in knowledge of human biases, deficiencies in substantive knowledge, and voids in representation from major groups of people who will likely be affected by the outcome of the decision. Be sure to include team members who do not have technical expertise in the decision domain–they will be some of the best people to look for value elements and blind spots in the decision process and to help reveal implicit assumptions held by different 'experts.' Ensure that all new team members have iterated through step number 2 before moving on to step number 4.

Step 4

Create a baseline of understanding regarding the decision making process and the specific decision domain by iterating through the high-level, key components of an outstanding approach to high-consequence, technologically advanced decisions shown in Figure 6.1. These include a sequence of four general activities:
1. Definition of the system bounding the decision domain with explicit reflection on what is known and what is uncertain about the system; the perpetual presence of 'unknown unknowns' should be explicitly acknowledged.
2. Conscious evaluation of critical thinking processes; at this point it is critical for all members of the decision team to review and understand what each of these critical thinking processes are and to practice identifying the application (or lack thereof) of these processes during discussions with team members–this approach will provide everyone on the team with the opportunity to put these processes into operation.
3. Assembly of the decision team with periodic reflection on whether representation of 'experts' and 'non-experts' is adequate versus the emerging collection of facts, values, and blind spots.
4. Review each of the risk perception/decision making biases/tendencies identified in the framework under the headings of normative knowledge, availability, and individual specific.

[51] See Winterfeldt and Edwards (1986) for a solid discussion of value trees, inference trees, and many other important tools that can improve high-consequence, multi-attribute decision making.

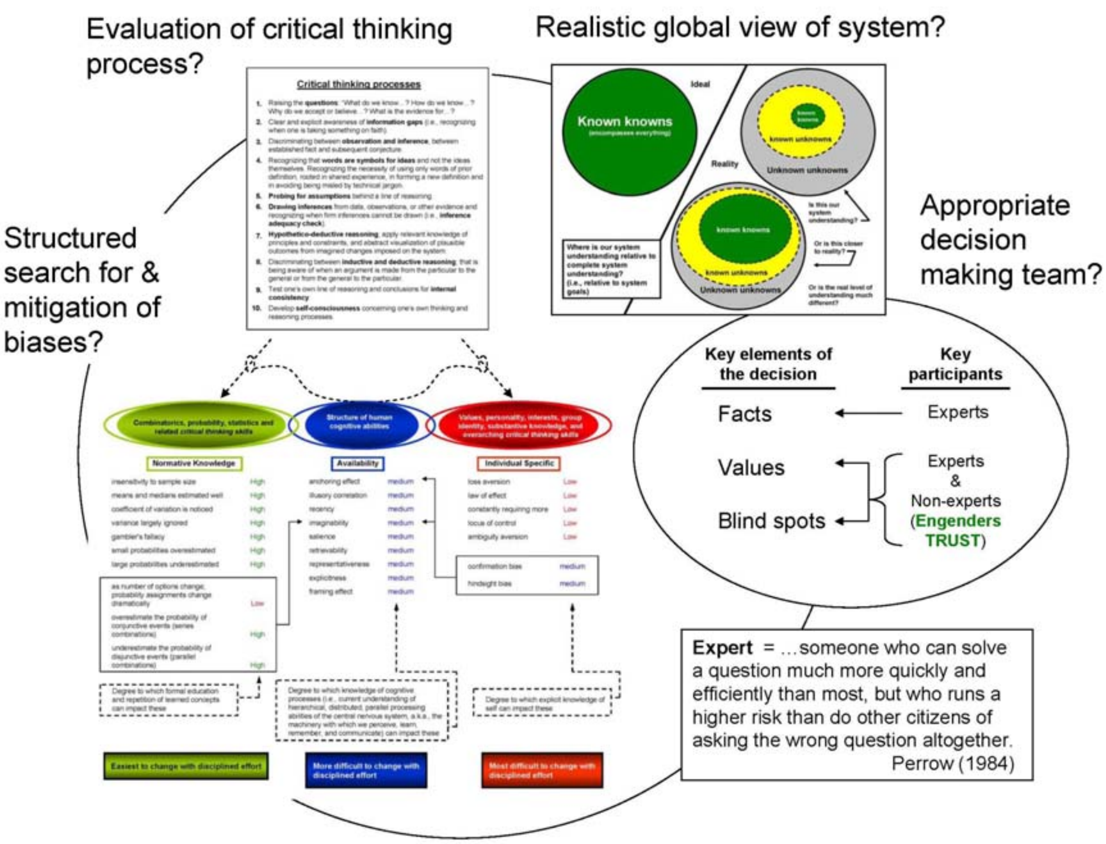

Figure 6.1. Overview of the key components of an outstanding approach to high-consequence, technologically advanced decisions.

<u>Step 5</u>

Repeat step number 2.

<u>Step 6</u>

Begin the detailed attempts to articulate the decision domain by gathering relevant data and information, determining what types of decisions could be made, selecting criteria for weighing decision options, and estimating the outcomes associated with decision options. These activities will require multiple iterations. Use a structured search during these iterations to identify and mitigate biases in the data and information. The bias framework should be applied using the following questioning strategy:

1. Does this ___ appear to involve normative knowledge, availability biases, or substantive knowledge?
2. Does this ___ appear to involve a specifically defined bias/tendency weakness?
3. Is there a suspected omission, misinterpretation, or lack of explicit interpretation that signals a deficiency in critical thinking?

Once outcome estimates and a decision plan begins to emerge, be sure to designate a facilitator to ask the decision makers to imagine themselves months or years into the

future with the knowledge that their plan has been carried out and that their plan has failed. This may be referred to as the time portal to failure (TPTF) technique.
Step number 6 is the heart of the decision making process and will require a considerable command of both the critical thinking processes and the framework of biases described in this report. Several helpful hints should be remembered while iterating through this critical step to help combat unhelpful decision making biases:

1. To combat confirmation bias, have a facilitator describe this powerful phenomenon (more than once) to the decision makers and those who will be affected by the decision. Openly admit that all members of the discussion must search for and question tendencies to disregard evidence that conflicts with current beliefs without well-founded evidence as measured using the ten critical thinking processes listed in this report. Normative suitability of information should be determined by normative experts on the decision team–but those experts must guard against overlooking arbitrary assumptions and highlight value judgments.
2. Do not automatically brand members of the public who raise concerns over decisions involving high consequence, advanced technologies as ignorant, ill-informed, or irrational. Always remember that people living in a technologically advanced society tend not so much to be in control of technology as dependent upon technology. There is a sense in which the increases in technical control have led to a decrease in social control. People must depend on (i.e., trust) specialists and specialized organizations they do not know and/or will not have the ability to control. This reality may make them attend more aggressively to anything that violates trust in the decision makers reflecting upon high-consequence, technologically advanced matters.
3. Remember that the availability categories of biases are an artifact of the way the 'machinery' of human cognition enables perception, learning, remembering, and communication. Explicit acknowledgement of this fact is a powerful way to depersonalize and diffuse conflicts of opinion and engender a disposition to seek mutual understanding.
4. Apply the ten critical thinking processes described in section three of this report to ensure that technical experts are focused upon **factual or technical questions**. Poor command of the critical thinking processes will prevent proper identification of authentically *factual* or *technical* questions.
5. Identify when aspects of a decision (model assumptions, metrics, etc.) involve answers to questions that are value judgments. Examples of such questions include: Is that safe enough? Is that level of uncertainty in the safety estimate acceptable? When it comes to questions of **values**, another word for 'scientist' is 'voter.' Ordinary citizens have just as much legitimacy in deciding questions of values as do scientists and engineers (Freudenburg 2001). Be wary of the occurrence of arbitrary decisions which are being passed off as 'scientific' when in reality they are merely arbitrary decisions made for convenience or due to lack of knowledge.
6. Apply the ten critical thinking processes, described in section three of this report, to aid in identifying **blind spots**. That is, items that have been overlooked. Additional terms for blind spots include *epistemic uncertainties* and unknown unknowns, which were discussed earlier in this report. Ordinary citizens are also very helpful for contributing to the discussion of blind spots as their thinking tends to be bounded much differently than those of a particular community of 'experts.' One should keep in mind Perrow's memorable definition of an *expert* as one who can solve a question

much more quickly and efficiently than most, but who runs a higher risk than do other citizens of asking the wrong questions altogether (Perrow 1984).

7. If scenarios are restricted to ≤ six steps, ≤ six decision metrics, etc., **stop** and **reflect;** it is highly likely that the decision makers are being confined by working memory limitations. Relieve the burdens on working memory by writing information down, use computer based tools, use record/replay equipment, etc. Recall the proper understanding of the principle of parsimony (a.k.a., Occam's Razor or the law of economy) is to interpret simplicity as the lack of unnecessarily improbable loose ends in a theory or model in favor of a theory or model that does a better "job" of tying things up. Stated differently: when deciding between two models that make *equivalent* predictions, choose the simpler one. Also recall the Pareto Principle, the analysis of the decision domain must aggressively seek understandings for many of the factors hiding in the 10–30 percent 'tail' of a Pareto tabulation of factors resulting in undesirable outcomes.
8. When people affected by a domain do not have substantive knowledge regarding the domain (i.e., they don't understand it), serious efforts at providing knowledge are required–that is, by incrementally building from simpler concepts to more challenging concepts in ways the audience can understand and using an inquiry based approach to check for increasing understanding (per Arons' critical thinking processes).
9. People are *loss averse* and seek to *gain*, often the concept of loss or gain is not associated with (or normalized by) total changes in wealth, safety, or anything else. Gain and loss are associated with the **mental accounts** in which money, safety, etc. have been apportioned. Find out what people really consider as gaining and losing in a particular decision domain, attempt to discover the mental accounts in which these concepts are binned, identify biases in these mental accounts which conflict with quantitative data, and make people aware of these biased mental accounts if they are present. This ties directly to framing. The general strategy for avoiding loss/gain, and framing effect decision traps is to deliberately frame a decision problem in more than one way. This means putting forth the effort to frame the presentation of data, information, decision criteria, and decision options in more than one way. Strive to achieve balanced presentations to avoid sending the decision team down a biased path and to maintain trust among those who will be affected by the decision.
10. People will start digging for representativeness at the expense of probability information–even when they are not competent to judge; all that is needed is for them to *believe* that they are a competent judge. To help mitigate this tendency, it can help to present a decision making team to non-experts (e.g., the general public) which is quite knowledgeable regarding many areas of the decision domain, but then to have the experts on that team publicly defer some technical areas of the decision to 'super-experts' in those domains. Publicly demonstrating deference from one expert to another will help make non-experts (e.g., the general public) feel that they are witnessing a truly rigorous scientific analysis and will encourage them to also reflect on parts of the decision where they may or may not be 'qualified' to contribute.
11. Remember that people consistently overestimate small probabilities and under estimate large probabilities. Provide numerous well-developed likelihood comparisons and explicitly state the underlying assumptions (to avoid weakening bonds of trust).

12. Throughout the decision making endeavor, actively search for instances where the hindsight bias is at work. Look for instances of "I knew that's how it was" and then probe thoroughly to see if an event of 're-writing history is taking place.' Take inspiration from Socrates who declared at his trial that "an unexamined life is not worth living" (Plato 1948 translation of work written before 347 B.C.).

A simplified summary of the above-mentioned six-step decision making process is provided in Figure 6.2. This overview can be used as the high-level guide for the decision making team as they carry out the decision process; however, frequent referral to the detailed exposition of the six-step approach is recommended to ensure effective application of the approach. Special attention should be paid to the twelve helpful hints in this section used to combat unhelpful decision making biases.

Six-Step Decision Making Approach

1. Explore individual specific attributes, reflect upon & review knowledge regarding cognitive capabilities, reflect upon & review normative knowledge. This may be conducted privately or between two or more decision makers.

2. Make an initial attempt at articulating the decision domain. Screen the initial articulation, initial data and initial information using the critical thinking skills and the unique framework for decision making biases/tendencies presented in this report.

3. After one or more iterations of step number 2, reflect upon the composition of the decision team. Add members to the team to ensure skill in normative knowledge, substantive knowledge, and adequate representation of 'non-experts.'

4. Create a solid baseline of understanding regarding the decision making process by iterating through the high-level, key components of an outstanding approach to high-consequence, technologically advanced decisions with new team members. That is, review system uncertainties, critical thinking processes, decision making biases/tendencies and again reflect upon the composition of the decision team.

5. Repeat step number 2.

6. Following iterations of decision domain articulation, designate a facilitator and conduct time portal to failure (TPTF) technique. Iterate one or more times through step 6 and refer often to the twelve helpful hints provided under the detailed discussion for step 6 in this report.

Figure 6.2. Simplified summary of the six-step decision making process.

The key question for the reader reflecting upon the process in this section might be: "If the decision maker or decision making team follows the process precisely, will it lead to better decisions?" The answer supported here is that following the process will lead to better decisions in nearly all cases, and the outcomes of those decisions will be, on average, superior to decisions made without a structured, bias-aware approach. That said, two things ought to be remembered: first, good decisions sometimes lead to bad outcomes and bad decisions sometimes lead to good outcomes; second, bad outcomes tend to be remembered much more easily than good outcomes–they are more *available*. Thus the prudent decision maker(s) must stay the course and tally the results over time to realize and appreciate the fruit of the highly-structured decision making approach when decisions regarding high-consequence, highly sophisticated technological systems are required.

Finally, it should be noted that a number of techniques have been developed for methodically weighing options such that bias effects may be mitigated. The author is not aware of any techniques which discuss/treat the range of biases and critical thinking processes mentioned in this report, but there are elements of those techniques that would benefit the diligent decision maker in organizing information, generating decision options, and evaluating potential outcomes of decisions. A helpful review of these techniques is provided in *Decision Analysis and Behavioral Research* by Winterfeldt and Edwards (1986).

7. Public Risks Related to Nuclear Power

As the economic and environmental impact arguments for increasing the use of nuclear energy for electricity generation and hydrogen production strengthen, it becomes important to better understand human biases, critical thinking skills, and individual specific characteristics that influence decisions regarding nuclear energy among the general public (e.g., trust of risk assessments, acceptance of new plants), and nuclear energy decisions made by high-level decision makers (e.g., energy policy makers & government regulators). To promote increased understanding and hopefully to improve decision making capacities, this section provides an example of *facts* and some aspects of the author's case for *value* in light of the previously presented decision-making material.[52] The overarching intent of this section is to provide an opportunity for the reader to begin applying the insights on risk perception and decision making presented earlier to a specific high-consequence domain involving technologically advanced systems. Pointers **[bracketed in bold text]** are planted in areas where biases/tendencies may be strong and to identify where previously presented concepts and techniques might be applied. The pointers are written as if a decision maker is reviewing and responding to the information in light of the biases/tendencies and critical thinking processes discussed previously. Also, the pointers are not intended to be complete nor exhaustive, but they are

[52] The particular perspective presented in this section regarding nuclear power for electricity generation was selected by the author to provide an example of how one may use the risk perception and decision making material previously described in this report to evaluate information as it might be presented for consideration by a decision maker. The views expressed in this section of the report do not represent an official position of Sandia National Laboratories on nuclear power.

suggestive of the appropriate reflections of the diligent decision maker when reviewing/organizing information related to a decision.

This example will present data supporting the claim that the risks from generating electricity with nuclear power are much lower than for several alternatives. It will also be proposed that one reason for heightened public fears regarding nuclear power is due to exaggerations by the media of risks and ‘mystery’ surrounding nuclear radiation. It may be true that such exaggerations by the media, which conflict with ‘nuclear power experts’ and government authorities such as the Nuclear Regulatory Commission breeds a skepticism that the ‘real truth’ about radiation risks is withheld, i.e., not trusting the stewards of nuclear power. This section will attempt to dispel the foundations of such skepticism by introducing the wealth of data and understandings regarding radiation in general and nuclear power production in particular. The credibility of nuclear power experts will also be defended by describing characteristics of the members of the National Academy of Sciences (i.e., those who are called in to provide authoritative technical positions on scientific matters).[53] The reader is encouraged to keep in mind the decision making framework, critical thinking skills, and decision making approach presented in the previous section to determine whether sufficient information is provided to make any decisions regarding nuclear power, and if so, to identify the specific decisions enabled. Again, pointers are provided in bracketed, bold font text.

The example begins immediately below and extends to the end of this section of the report:

Of key importance when discussing the topic of risk associated with any type of behavior or phenomenon is to quantify the relevant risks based upon likelihood of occurrence and consequence severity. It is virtually useless to place qualitative assignments on risks without considering quantitative factors. As stated by Blaise Pascal in the 17th Century, “Fear of harm ought to be proportional not merely to the gravity of the harm, but also to the probability of the event” (Hacking 1975) **[addressing potential deficits of reader in substantive and normative knowledge]**. For example, someone may say that there is a possibility that a single high-energy neutron may disturb an atom within a molecule inside a human cell that leads to a cancerous growth in one person that eventually leads to severe harm or death; this qualitative realization may lead to the decision that all substances that can potentially release a high-energy neutron (or other high-energy particles) should be avoided at *all costs.* As a second example, it is conceivable that the ignition of a small amount of gasoline vapor could lead to a fire that could destroy an entire city and cause hundreds of thousands of fatalities; this qualitative realization might lead to the decision to ban the production and use of all petroleum-based fuels.[54] Accepting either of the two decisions above would be foolish when quantitative costs versus benefit data are considered **[keep in mind availability biases and loss vs. gain sensitivity]**.

[53] Another aspect of this lack of trust of nuclear experts is fueled by the minority of anti-nuclear scientists and engineers which often receive a disproportionately large share of media coverage. J. Samuel Walker carefully reviews such phenomena in his accurate and engaging historical review of nuclear power in the U.S. and detailed account of the accident at Three Mile Island, Unit 2 (Walker 2004).
[54] Another partial solution might be to store gasoline in ridiculously small individual containers.

The risks related to nuclear power to be considered in this section include exposure of the public to radiation from nuclear materials during normal operation of nuclear power facilities, or following an accident (nuclear power plant core melt, damage to a nuclear material container, etc.). The topic of risks related to nuclear power can be examined using quantitative approaches since a great deal is known about the effects of radiation sources (i.e., X-rays, Gamma rays, neutrons, electrons, alpha particles, and various other fission products) versus other environmental agents **[trying to shore up substantive knowledge deficits in the reader]**. Tools are also available to accurately measure radiation levels. The biological effects of other hazardous materials such as sulfur dioxide, nitrogen oxides, hydrocarbons, volatile organic compounds, and toxic metals (e.g., cadmium, arsenic, nickel, chromium, beryllium, etc.) are much less understood **[can I trust this strong statement?]**. Radiation dose-response relationships have been studied more thoroughly and are better understood than dose-response relationships for many chemical and biological agents (Voelz 1975; Saunders and Wade 1983; Cohen 1990). It is noteworthy that public perceptions of risks related to the use of nuclear power to generate electricity seem to be much higher than perceived risks related to other electricity generation methods.[55]

Radiation sources may be perceived by some in the general public as a new phenomenon with largely unknown associated hazards. This view is patently incorrect since radiation and radioactivity was discovered near the end of the nineteenth century, and many subsequent applications of radiation have generated voluminous data regarding radiation effects in animals and humans. The body of knowledge regarding radiation hazards has developed from measured exposures and effects in the following groups: radiographers in the late 1800s and early 1900s; underground metalliferous miners; radium luminizers;[56] medical treatment[57] of ankylosing spondylitis (i.e., fusing of joints–primarily in the spine), ringworm, and cancerous tumors; diagnostic X-rays; accidents during the Manhattan project and subsequent operation of the nuclear weapons complex (both the US and UK weapons programs); Hiroshima and Nagasaki survivors; Marshall Islanders exposed to radioactive fallout from nuclear weapon tests; occupational exposures in the nuclear power industry and those impacted by the accident at the Chernobyl Unit 4 reactor in Ukraine (Saunders and Wade 1983; Knief 1992).

In order to develop an argument defending the knowledge base regarding human health effects due to radiation exposure, it is essential to describe radiation, how radiation may enter the body, and what it might do inside the body. First, radiation is a term that can generically be used to describe a range of both ionizing and non-ionizing radiation. Ionizing radiation is itself a relatively generic term which refers to directly ionizing

[55] It is quite possible that while much is known about radiation and its effects on the human body, relatively little of this knowledge has been presented to the general public in a clear and direct matter using non-technical terms.

[56] At one time, watch dials were painted with beads of radium to provide visibility during darkened conditions. Unfortunately, the women who worked in the watch factories maintained a fine point on their paint bushes by licking the tips. This practice caused accumulation of radium (a decay product of either U-238 or Thorium-232 and is an alpha particle emitter) near the thyroid glands, which often led to large, eventually fatal tumors.

[57] This includes study of both patients and radiological health workers.

radiation such as charged particles (e.g., protons, alpha particles, beta particles, etc.), and indirectly ionizing particles without an electric charge (e.g., neutrons, X-rays and gamma rays). Both of these types of ionizing radiation can displace electrons, leaving a trail of ion pairs, i.e., electrons or charged atoms or ions (Saunders and Wade 1983). This type of radiation can be contrasted with non-ionizing radiation; for example, microwave radiation which does not ionize (i.e., remove electrons from atoms; thus, it is too weak to cause direct chemical reaction in a significant manner). Yet such radiation can still cause damage as it adds to the collective thermal energy of a body. When discussing radiation associated with nuclear power or nuclear weapons, the topic is often understood to be focused on ionizing radiation.[58]

Ionizing radiation, which causes the displacement of orbital electrons in the atoms through which it passes, can damage tissues in the human body by causing immediate biochemical changes that may eventually lead to chromosomal aberrations, cell death, malignant transformation, or cellular repair **[addressing question: why is it bad?]**. Simply put, ionization produces free radicals or ionized molecules which react chemically with adjacent atoms and molecules to alter cell physiology (Voelz 1975). For example, in living tissue, the 70% or more water content of the body is a major target of these reactions in the following way: an H_2O molecule can be hit by proton, alpha particle, or gamma radiation causing the liberation of an electron which then leads the unstable cation of H_2O to break down into H^+ (a proton) and an OH (hydroxyl) radical. Both of these 'free radicals' may initiate a host of undesirable chemical reactions depending on the particular location of the reaction(s) and whether biologically important molecules (e.g., DNA, proteins, etc.) are nearby (Voelz 1975; Brady and Holum 1984; Knief 1992). Furthermore, cell damage due to radiation tends to be greatest in cells with the highest degree of differentiation or are multiplying most rapidly, which leads to the following descending hierarchy of susceptibility: lymph; blood; bone; nerve; brain; and muscle cells). Table 7.1 below shows some of the deleterious effects of such radiation induced damage observed in a range of mammals (i.e., both in human and non-human populations).

Table 7.1 Selected examples of mammalian radiation induced biologic damage that can occur. Adapted from Voelz (1975).

Level of Damage	Important Effects
Molecular	Damage to DNA, RNA, enzymes, interferes with metabolic pathways
Subcellular	Damage to cell membranes, mitochondria, lysosomes, nucleus, & chromosomes
Cellular	Inhibits cell division, kills cells; causes carcinogenesis
Organ	Injures bone marrow, intestinal tract, cardiovascular and CNS, with minor changes to acute deaths in severe exposures
Whole Body	Death, cancer, gene & chromosomal mutations – alter offspring genes

[58] However, the intense ionizing radiation in a nuclear burst heats the atmosphere, which in turn re-emits energy that finally does it damage as non-ionizing energy in the form of intense light and a thermal wave that can start fires and produce severe burns. Ultraviolet light, which is responsible for sunburn and is present in a nuclear burst, is essentially energy in the range between strictly ionizing and strictly non-ionizing.

The different types of ionizing radiation have some very different characteristics. The ionization process for charged particles such as alpha particles, protons and beta particles when they enter tissue occurs quickly as they will either directly collide with an oppositely charged particle or act as an electrical repulsive or attractive force to 'rip' away an orbital electron **[what types of ranges do these charged particles have in biological tissue–that is the information I want?]**. In contrast, electromagnetic radiations such as x-rays or gamma rays behave somewhat differently. These radiations are composed of a stream of tiny electromagnetic bundles of energy called photons that do not carry an electrical charge. Therefore, they easily pass through open spaces in material until photons interact with an atomic nucleus or an electron. In other words, electromagnetic radiation can penetrate much deeper into material. Neutrons, also without charge, may travel long distances **[how far?]** through material before collision (and usually absorption) into another nucleus **[continuing to shore up substantive knowledge, but there may be some items missing here]**.

In addition to whether radiation consists of charged or uncharged entities, it is critical to have some relative understanding of the energy levels and penetrating power associated with each type. For charged particles, the range at which ionizations will occur in a given material is nearly directly proportional to the energy level. For example, a 5 million electron volt (MeV) alpha particle will penetrate roughly 3.5 cm in standard air and about 35 millionths of a meter in water or soft tissue **[starting to answer 'how far?' question]**. An 8 MeV alpha particle has roughly double the range in both materials. The penetrating distance of electromagnetic radiation (having no charge) in similar low density materials is exponential **[exponential? I vaguely remember what this represents, must review normative knowledge. Is it always exponential?]** with distance. In order to approach proportional stopping power for x-rays or gamma rays, a shield of very dense material (e.g., lead) needs to be used (Voelz 1975). Table 7.2 summarizes some characteristics of ionizing radiation and Figure 7.1 below shows a rough comparison of the relative penetrating power different radiation types.

Table 7.2 Radiation characteristics. Taken from Voelz (1975).

Name	Symbol	Character	Mass[59]	Charge	Source examples
X-rays	x	Electromagnetic energy	0	0	X-ray tube
Gamma	γ	Electromagnetic energy **[are X-rays and Gamma rays simply different points on a continuum? Do they originate in the same manner?]**	0	0	Co-60, Ir-192
Alpha	α	Particulate, helium nucleus	4	++	Pu-239, Po-212
Beta	β	Particulate, electron	1/2000	-	Sr-90, H-3
Neutrons	n	Nucleus particle	1	0	U-235 fission

[59] The atomic mass unit (AMU) is used here so that a neutral carbon-12 atom has a relative mass exactly equal to 12. This is equal to ~ 1 AMU for both a proton and a neutron.

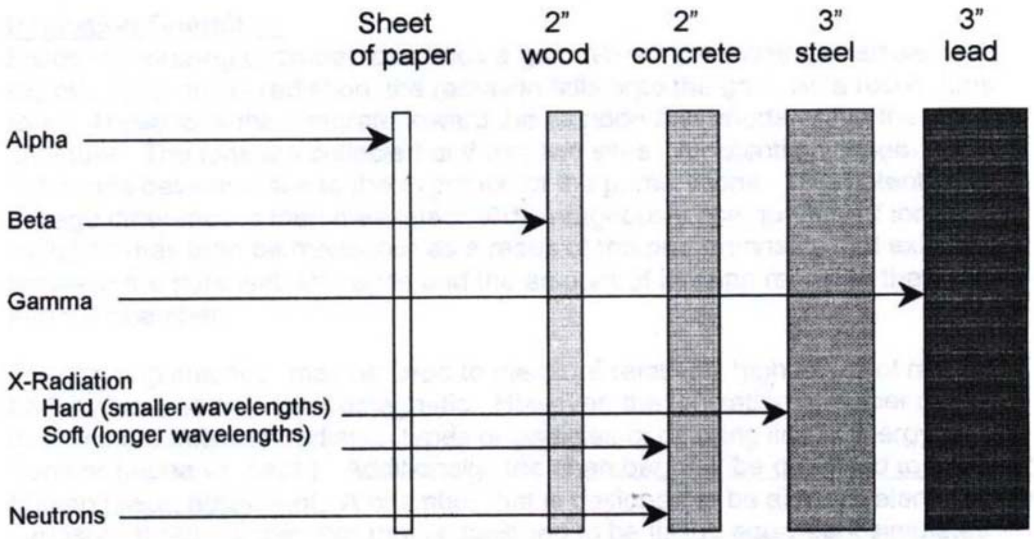

Figure 7.1. Approximate relative penetrating power for different types of ionizing radiation. Taken from Talty (1988, p. 633) **[What energy levels are represented in this figure? Does the penetrating power of a neutron depend upon the mass of the nucleus of the atoms in the material? (e.g., efficient transfer of kinetic energy to light nuclei)]**.

Another convenient measure of the macroscopic effect of radiation versus target material interactions (in addition to the range of penetration) is referred to as the linear energy transfer (LET). The LET is the energy deposition of a charged particle or high energy photon per unit distance traveled. Particles with larger masses and more charge will have a higher LET and electromagnetic radiation will have a much lower LET **[careful definition of words and acronyms; critical thinking process #4]**. When LET values are high, then within the given target area there will be many ionization events with a high likelihood of causing harmful biological effects, even at relatively low doses. Conversely, low relative doses of low-LET radiation **[how low?]** will generally produce isolated **[how isolated?]** biological effects which may be mitigated via the organisms molecular repair processes (Saunders and Wade 1983). Table 7.3 shows the relative order of magnitude ranges and LET for three energy types, and Figures 7.2 and 7.3 show qualitative contrasts between high and low LET radiation.

Table 7.3 Order of magnitude comparison of range and LET for naturally occurring radiations in a specified material. Taken from Knief (1992).

Radiation	Relative range	Relative linear energy transfer [LET]
Alpha	1	10,000
Beta	100	100
Gamma	10,000	1

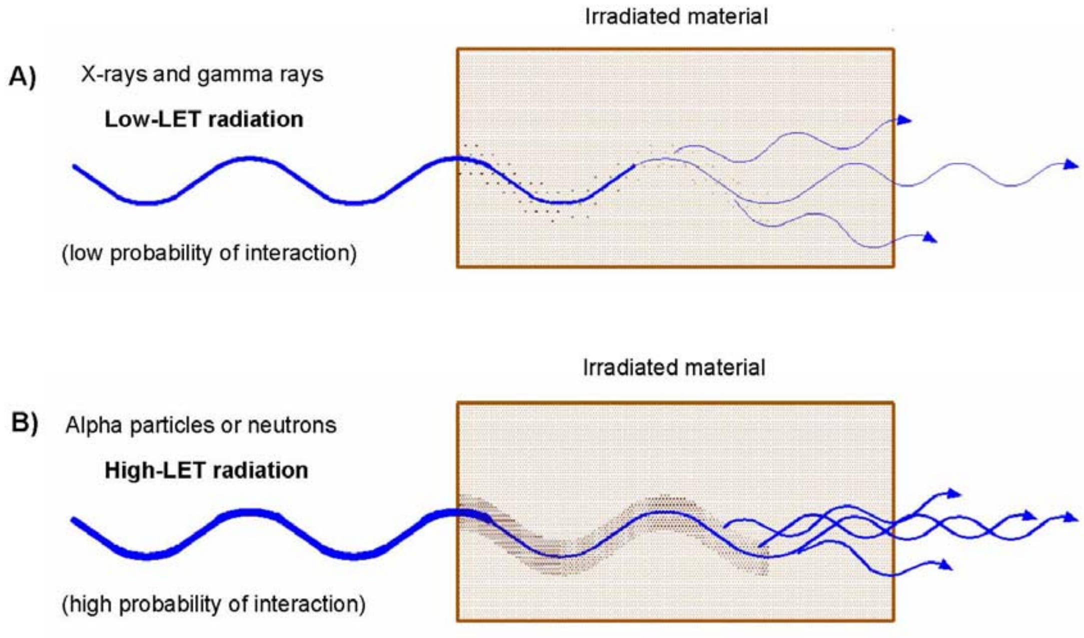

Figure 7.2. Qualitative contrast between the degree of interaction between low (A) and high (B) LET radiation with nearby materials. Adapted from Saunders and Wade (1983) **[Is this visualization accurate?]**.

[Visualization is very helpful for addressing substantive knowledge deficiencies in this topic area, but I wonder if I need to know these details?]

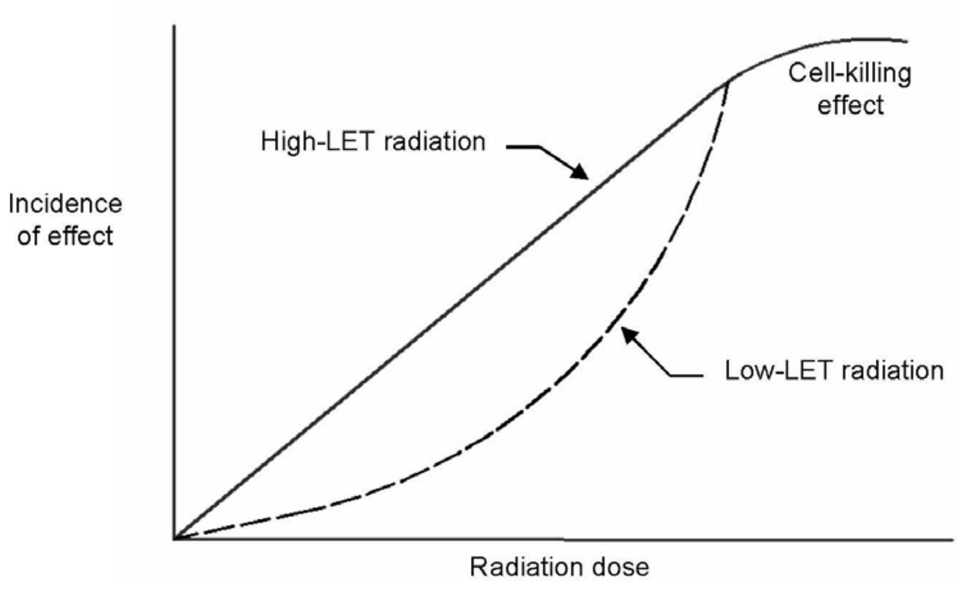

Figure 7.3. Qualitative contrast between low and high LET radiation with respect to biological cell-killing effect. Adapted from Saunders and Wade (1983). **[Do I need to have a sense for the units here? Is a qualitative representation sufficient?]**.

The different radiation types introduced above not only have different material penetrating powers, but they also have other characteristics which determine the types of damage they may cause to the human body. Alpha particles, although unable to penetrate skin, pose a danger to humans if they are introduced into the body such as via ingestion of an alpha emitting radionuclide where they are able to ionize atoms and molecules (i.e., produce free radicals) very close to critical biological tissues **[this section is helping to answer question: How and why is it dangerous?]** . Beta particles have a greater ability to penetrate skin, but their reduced LET gives them much less ionization power than alpha particles, which means that they are not often of great concern unless a beta emitting radionuclide gets inside the body **[how do beta emitters get in the body?]**. Gamma radiation sources can produce two main types of harmful exposure. The first is due to a large gamma source that produces radiation which penetrates the whole body. The second is ingestion or introduction of gamma emitting radionuclides into the body which emits some gamma radiation, but more importantly, the gamma rays may cause the production of alpha or beta particles **[how does this occur exactly? Need to get more information]** that then become the more intense ionization source for nearby tissues.

Neutrons, which are emitted from fissioning radionuclides or as a result of bombardment of nuclei with high energy particles (e.g., protons in a particle accelerator) may occur at various energy levels. The usual categories of neutron energies include thermal (~0.025eV), epithermal or slow (up to 100 eV), intermediate (100 eV–100 keV), and fast neutrons (>100 keV) **[why is this important?]**. The neutrons that people might come in contact with as a result of fission reactor operations or particle acceleration would likely be fast neutrons, which can cause very significant damage to the body (Voelz 1975) **[that's why this is important]**.

At this point, a logical question might be, "Given that various types of radiation sources have different capacities for producing harm in human, especially when entered into the body, how exactly does a radiation source enter the body?" Figure 7.4 below displays the various avenues for radiation exposure to the body. Other than direct irradiation with penetration through the skin, radionuclides may enter via the lungs (e.g., inhaled radon gas), through food stuffs, etc.[60] **[this information makes me feel better about the possibility of controlling exposures–and communicating to others how control can be accomplished].**

---

[60] The many subtleties of radiation exposure related to specific radiation sources and how they tend to cause damage in humans is outside the scope of this report. The important point to note is that extensive study regarding these subtleties has been and continues to be carried out within a very well established health physics community. See Voeltz (1975), Saunders and Wade (1983), and Knief (1992) for more detailed overviews and references to health physics texts and primary research studies **[lot's of appropriate observation & inference, and induction & deduction alluded to here–good, but need to research more]**.

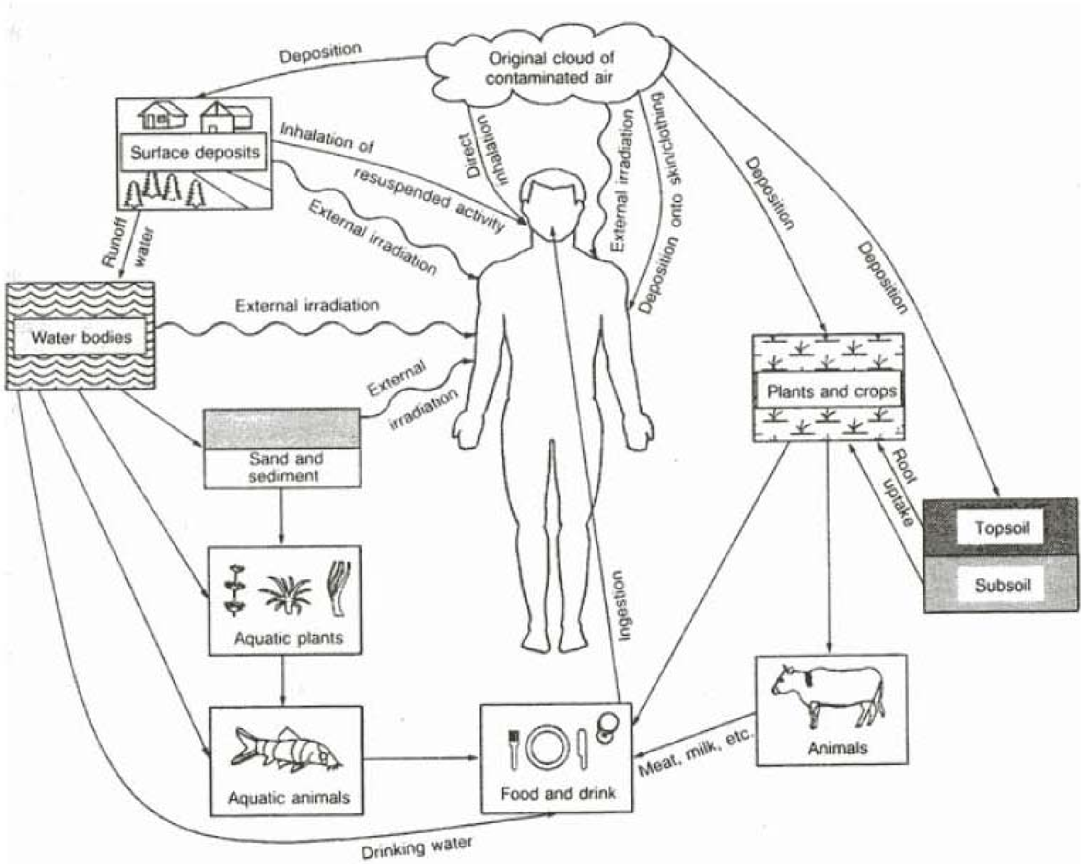

Figure 7.4. Environmental pathways of human radiation exposure. Taken from IAEA (1991) **[What about cosmic rays? Those were mentioned earlier, is there an omission here?]**.

The probability of any one highly-energetic particle causing cancer in humans has been estimated to be 1 chance in 30 x 10^{15} **[How exactly was this estimated? Should I take this on faith? Must dig deeper; where does the linear versus threshold model enter in here–I think I've heard something about that before.]**, and it is also known that each person is bombarded by such radiation particles (on average) at the rate of approximately 15,000 per second (i.e., every second of one's life) due to background radiation (including cosmic and terrestrial sources). In fact, a common medical X-ray produces over one trillion such particles into the human body as a consequence of X-ray absorption (Cohen 1990). Returning to the example of the gasoline initiated fire, the likelihood that the circumstances would converge for a city-destroying fire are arguably much higher than 1 in 30 x 10^{15} **[how likely would something like that be?]** . Yet it is still not reasonable to eliminate all petroleum fuel products based on the qualitative realization that such a disaster is theoretically possible.

Considering again the example of the tremendously low probability for health effects due to being hit by a single high-energy particle, it is important to realize that the nature of this risk is similar to many others faced in daily life. A multitude of chemical, physical, and biological processes threaten a person's existence with much higher likelihood than those related to even very severe nuclear radiation releases. Reasonable people recognize that life is inherently risky. For example, one study has estimated that there is an equivalent risk (defined as increased risk of death within one year by 1 in 1 million) shared by each of the following: smoking 1.4 cigarettes; riding 10 miles by bicycle; eating 40 tablespoons of peanut butter, drinking 30 12-ounce cans of saccharine sweetened diet soda; and living within 5 miles of a nuclear reactor for 50 years (Inhaber 2004) **[what are the normative data supporting these inferences? Probably averages of data from many epidemiological studies–must investigate]**.

Air pollution, a byproduct of fossil fuel burning and other industrial processes, is also known to increase the risk of death. A Harvard University research study sponsored by the US DOE concluded that air pollution is likely causing 100,000 deaths per year due to heart and lung disease and another 1,000 per year due to cancer (Ozkaynak and Spengler 1985; Cohen 1990). The central point in the above discussion is that in order to make a rational decision regarding the risks of a behavior or activity, it is essential to review quantitative data regarding the likelihood and severity of injury/damage associated with that risk. In the following paragraphs, it will be shown that the types of radiation exposures resulting from the nuclear power industry are extremely low. Additionally, it is proposed that heightened public fears **[How heightened are they relative to fears about risks from other energy sources?]** due to radiation exposure are caused in large part by biased media attention.

We now present a brief discussion of the units used for measuring radiation doses from Saunders and Wade (1983). The curie (Ci) is a measure of the activity of a radioactive source; 1 curie = 1 Ci = 3.7×10^{10} disintegrations per second (i.e., the number of nuclei that 'disintegrate,' which means spontaneous transformation of the number or the internal arrangement of protons or neutrons in the nucleus of a radionuclide). It is important to note that this measurement unit says nothing about the specific type of decay activity. An example of using this unit is the following: "The activity of spent fuel rod # 1275 on November 1, 2004, was 8.3×10^4 Ci." The half-lives of the nuclides that comprise the fuel and the types of radiations they emit have no bearing **[is this really true? It has no bearing?]**. The SI unit of source radioactivity is the Becquerel (Bq), where 1 Bq equals 1 disintegration per second. The roentgen (R) is a measure of *exposure*, which is the ability of a beam of x-rays or gamma-rays to deliver energy to a material through which they pass. One röentgen is the exposure that would deliver 8.78 mJ to 1 kg of dry air at standard atmospheric conditions. For example, "This dental x-ray beam generates an exposure of 400 mR/s." The rad, which is an acronym standing for Radiation Absorbed Dose, is the measure of the dose (i.e., energy) actually absorbed by a specific object. An object, which might include a person (i.e., whole body) or a body part (e.g., the hands), is said to have received an *absorbed dose* of one rad when 10 mJ/kg have been delivered to it by ionizing radiation.[61] An example of using this unit is the following: "A whole-body short term gamma ray dose of 400 rad will cause death in 50% of the population exposed within 60 days without medical treatment." The SI unit of absorbed dose is the Gray

[61] Ionizing radiation is a relatively generic term which refers to directly ionizing radiation such as charged particles, and indirectly ionizing particles without an electric charge such as neutrons, X-rays, and gamma rays. Both of these types of radiation can displace electrons, leaving a trail of ion pairs, i.e., electrons or charged atoms or ions (Saunders and Wade 1983). This type of radiation can be contrasted with, for example, microwave radiation which does not ionize (i.e., remove electrons from atoms).

(Gy), where 1 Gy equals 1 J/kg of transferred energy **[this section is attempting to increase substantive knowledge by defining terms which can then be used to arrive at common understandings of risks of radiation–good approach; I'm learning a lot]**.

Next, is the rem, which is an acronym for Röntgen Equivalent in Man and is a measure of *dose equivalent*. This unit takes into account the fact that even though different types of radiation (e.g., gamma rays, neutrons, x-rays, etc.) may deliver the same energy per unit of mass to the human body, they do not have the same biological effect. The dose equivalent measured in rems is found by multiplying the absorbed dose in rads by a *Relative Biological Effectiveness* (RBE) factor. Examples of the RBE include: x-rays and electrons → RBE ~ 1, slow neutrons → RBE ~ 5, etc. An example of using this unit is the following (Saunders and Wade 1983; Halliday and Resnick 1988): "The National Council on Radiation Protection recommends that no individual who is (nonoccupationally) exposed to radiations should receive a dose equivalent greater than 500 mrem (= 0.5 rem) in any one year." The SI unit of dose equivalent is the sievert (Sv), where 1 Sv = (1 Gy)(1 Q) = 100 rem. Figure 7.5 and Table 7.4 below provide additional description and a summary of radiation units **[words are symbols for ideas; I need to memorize these foundational ideas to prepare for the decision]**.

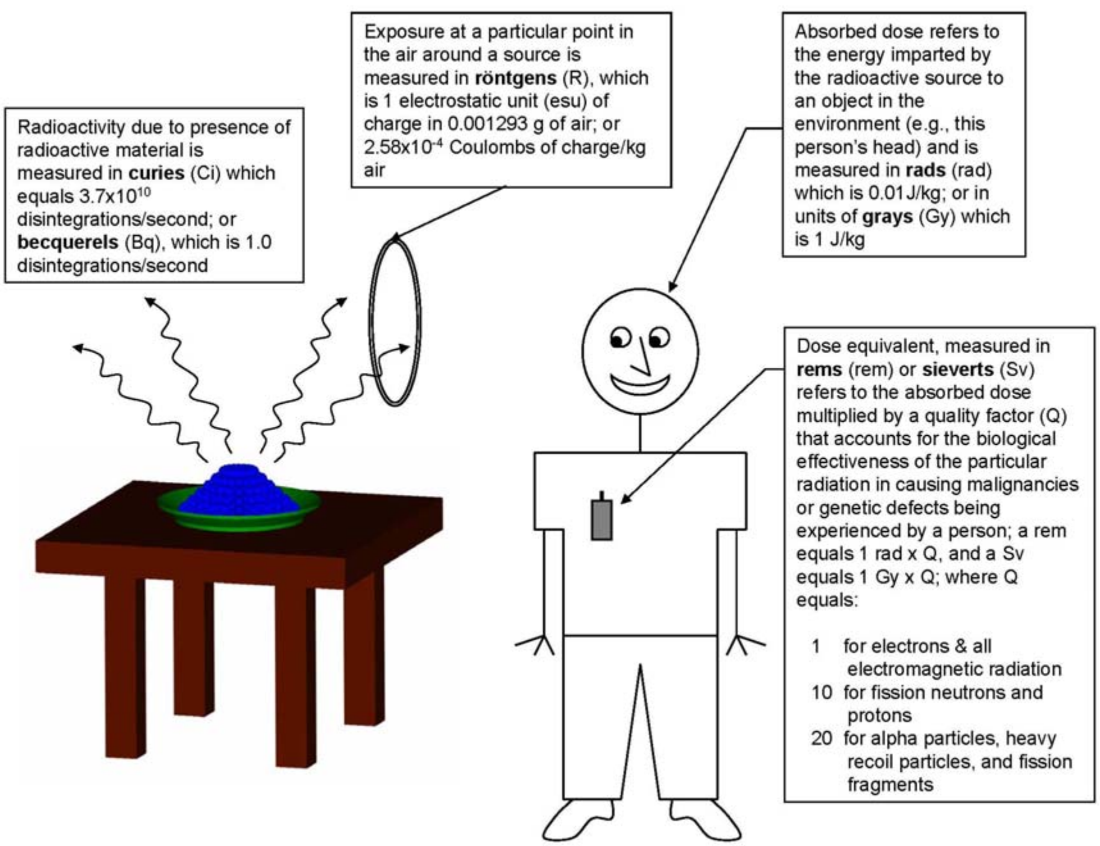

Figure 7.5. Common radiation units and a visualization to help explain what the units actually represent. The person in the figure is shown wearing a dosimeter badge. Adapted from Halliday and Resnick (1988) and Saunders and Wade (1983).

Table 7.4 Summary of radiation units. Taken from Saunders & Wade (1983).

Quantity	Name	Abbreviation	Units
Radiation exposure	roentgen	R	2.58 x 10^{-4} Coulombs /kg air **[what is a Coulomb? Must check]**
Radioactivity	curie becquerel*	Ci Bq	3.7 x 10^{10} disintegrations/sec 1.0 disintegrations/sec
Absorbed dose	rad gray*	rad Gy	0.01 J/kg 1 J/kg (= 100 rad)
Dose equivalent	rem sievert*	rem Sv	Rad x $Q^{\ddagger}$ Gy x Q (= 100 rem)

* SI units; ‡ See text and Figure 7.5 for an explanation of Q.

The most common unit of radiation exposure is the millirem (mrem). A single mrem is equivalent to being struck by roughly 7 billion radiation particles, given the differences in health effects due to different particle types. Low-level radiation is typically considered to be a dose below 10,000 mrem **[Is this dose level correct? Is this meant to be a dose equivalent over a year or some other time period? Need to check…]** (Cohen 1990). It has been estimated that persons in the United States are exposed to the following *average* levels of radiation per year: 30 mrem from cosmic rays; 20 mrem from naturally occurring radioactive materials in the ground (uranium, potassium, thorium, etc.); 10 mrem from building materials (brick, stone, plaster, etc.); 25 mrem from radiation internal to the body (i.e., mostly potassium-40); and an average of 80 mrem from dental and other medical X-rays (NCRP 1975; Cohen 1990). The sum of these radiation exposures reveals an average annual exposure of 165 mrem. In addition, there is an average annual radon exposure level of 200 mrem in the US for a total of ~365 mrem overall per year. Background radiation levels for specific locations vary widely depending upon such factors as geographical altitude **[Does this mean that radiation exposure is greater when I travel in an airplane? How much greater?…I fly a lot]** and nearby geologic formations.

The levels of radiation exposure resulting from highly publicized ‘radiation releases’ include the following: Three Mile Island (TMI) accident–average exposures in surrounding areas = 1.2 mrem; highest exposure from a leak of radioactivity from a low-level waste site in Kentucky = 0.1 mrem; highest exposure from a leak at a nuclear power plant in New York = 0.3 mrem. Inflammatory headlines from these events included statements such as “RADIATION,” “It’s Spilling All Over the US,” and “There’s No Place to Hide” (Cohen 1990) **[the availability biases are at work here]**. These facts raise a couple of questions: first, should there be concern over any amount of radiation over the yearly averages stated previously? Second, were there greater ‘accident potentials’ indicated by previous accidents that justified the concerns that were raised at the time?

The question of whether any ‘above natural’ radiation exposure should be tolerated **[this is a ‘value’ question]** can be addressed succinctly by looking closer at the US average radiation exposure values presented in this section. As with many ‘average’ statistics, there are wide variations enclosed within the underlying distribution. For example, people in the Rocky Mountain states (Colorado, Wyoming, New Mexico, & Utah) have twice the average radiation exposure (i.e., ~330 mrem, not including radon) due to higher proportions of uranium in the soil and higher elevations that reduce the amount of protective atmosphere from cosmic rays. Yet the cancer rate in Colorado is 35% below the national average and the incidence of leukemia, the cancer most strongly associated with radiation exposure, is 14% below the national average. Wyoming, New Mexico, and Utah have leukemia rates 39% below the national average. This reinforces the claim **[reinforces yes, but does not prove; this is a strong inference]** that radiation is not one of the important causes of cancer (Cohen 1990). By far the most significant and dangerous source of radiation comes from radon gas. This gas is found in many homes and when breathed in it irradiates the lungs. The average lung exposure due to radon is equivalent to receiving 200 mrem to all organs in the body. In some states the average exposure is 600 mrem, and over 12 million Americans get exposed to more than 1,000 mrem per year. Ironically, the residents of Pennsylvania near the TMI plant are exposed to more radiation per day from radon than they received from the reactor accident in 1979[62] **[What are the estimates regarding the significance of radon as a cause of cancer? When radiation exposure from radon is compared to the TMI accident, is that a time period of one day, one week, the first year after the accident…does that matter?]** (Cohen 1990).

To put these levels of exposure into perspective, it is helpful to note that in most situations the risk of dying from cancer is increased by roughly 1 chance in 4 million for every mrem a person receives. This risk conclusion has been reached independently by both the U.S. National Academy of Sciences Committee on Biological Effects of Ionizing Radiation and the United Nations Scientific Committee on Effects of Atomic Radiation (UNSC 1988; Cohen 1990; USNAS 1990)[63] **[this is addressing the ‘how do we know?’ question]**. This increased risk equates to a reduction in life expectancy of 2 minutes, which is equivalent to the following (Cohen 1990) **[Are these really ‘equivalent’, how were they calculated?]**:

- Crossing the street five times (i.e., from the average probability of being killed while crossing a street)
- Taking a few puffs on a cigarette (i.e., each cigarette smoked by a habitual smoker reduces life expectancy by 10 minutes)
- An overweight person eating 20 extra calories (e.g., a quarter of a slice of bread and butter)
- Driving an extra 5 miles in an automobile

[62] The reader is encouraged to keep in mind that the units of comparison for these different types of radiation exposure are in rems; i.e., adjustments for different types of biological impact have been included.
[63] This claim is supported by extensive data gathered from Japanese atomic bomb survivors, and medical treatment studies in England, Germany, Canada, Denmark, Finland, Norway, Sweden, United States, Potugal, Nova Scotia and Yugoslavia as well as studies on miners around the world and numerous other studies. For details, see Saunders and Wade (1983) and Cohen (1990).

In addition to reviewing actual risks from levels of radiation exposure, it is also helpful to investigate the role of the media in generating public concern for radiation exposure and to investigate whether having a heightened fear of a serious radiation accident due to the nuclear power industry is a reasonable position. Bernard Cohen, author of *The Nuclear Energy Option* conducted a search of entries in the New York Times Information Bank on different types of accidents and compared them with how many fatalities resulted from each type in the U.S. This search was conducted for the time period between 1974–1978, which was before the increase of fears after the TMI accident in 1979. Cohen found that there were an average of 120 entries per year regarding motor vehicle accidents (MVA), which kill ~50,000 people per year; 50 entries per year on industrial accidents, which kill ~12,000; and 20 entries per year on asphyxiation accidents, which kill ~4,500. These numbers reveal 1 entry per 417 MVA deaths, 1 entry per 240 industrial accident deaths, and 1 entry per 225 asphyxiation accident deaths. In stark contrast, there were 200 entries per year for accidents involving radiation, yet not a single fatality had been caused by radiation for over a decade (Cohen 1990) **[this is trying to quantify the occurrence of cases where powerful availability biases are present; these may/or may not have been consciously manipulated–very important]**.

It is quite apparent that media coverage tends to exaggerate the radiation hazards from past accidents relative to the minimal damages incurred. It is also proposed that general media sources scare the public by presenting radiation as a new and mysterious form of damaging energy. In reality, natural radiation sources have always been bombarding mankind. Furthermore, a human cell cannot tell the difference between a high-energy particle from interstellar space or one that originated from plutonium-239 in a manmade reactor **[this presentation of radiation as 'new' probably has impacted development of individual specific biases in opposition to nuclear power/radiation]**. Again it ought to be reiterated that the health effects of radiation are much better understood than those of air pollution, additives in food, chemical pollutants in water, or nearly every other environmental agent (Cohen 1990). This greater understanding is due to radiation having well-understood mechanisms for interacting with matter. Other environmental agents may have hundreds of important components that interact in complex and poorly understood ways. As stated by Cohen (1990):

> Radiation is easy to measure and quantify, with relatively cheap and reliable instruments providing highly sensitive and accurate data, whereas instruments for measuring other environmental agents are generally rather expensive, often erratic in behavior, and relatively insensitive. And finally, our knowledge of radiation health effects benefits from a $2 billion research effort extending over 50 years. More important than the total amount of money is the fact that research funding for radiation health effects has been fairly stable, thereby attracting good scientists to the field, allowing several successive generations of graduate students to be trained and excellent laboratory facilities to be developed **[With the recent advances in analyzing chemicals since 1990, does this type of statement still hold? Of course, new sensing equipment would not have been around long ago to help with retrospective, epidemiological studies – should I dig deeper into the assumptions here?]**.

Another observation by Bernard Cohen that is important to mention here is his personal impression of how the major TV media organizations have perceived problems related to nuclear power. He feels that the "TV people" regard the official committees of scientific experts to be pawns of the nuclear power industry rather than objective authorities **[this highlights the difficulties with advanced technologies]**. Often, the experts featured on television are members of the National Academy of Sciences. Cohen responds to such criticisms by noting the following points regarding the objectiveness of this elite group[64] **[trying to address substantive knowledge deficits and engender 'trust' of experts]**:

- The National Academy of Sciences is a nonprofit organization chartered by the U.S. Congress in 1863 to further knowledge and advise the government. It is comprised of approximately 1,000 distinguished scientists from all branches of science. A majority of these scientists being tenured, university professors, which is among the most secure types of employment in the world **[Do these distinguished scientists ever get exposed to intense political pressure that can be used to threaten funding sources to maintain particular research programs?]**.
- Believing that such reputable scientists would conspire to deceive the public is absurd, if for no other reason than that it would be easy to unveil what they had done, which would immediately devastate their professional standing **[How transparent are their analyses? Can the assumptions be readily discerned?]**.
- These are people who have chosen a field dedicated to the protection of the health of others **[Should we infer that this actually applies to all NAS members? Or does it apply only to a subset of members?]**; with respect to radiation physics experts, many of them are medical doctors who have turned down lucrative careers in medical practice to become research scientists. "To believe that nearly all of these scientists were somehow involved in a sinister plot to deceive the public indeed challenges the imagination" (Cohen 1990).

As for whether it is reasonable to have heightened fears regarding a serious accident within the nuclear power industry, we must again turn to quantitative estimates of the likelihood and severity of a major accident in the United States. The probability of having a core meltdown accident, that is, one in which there is both damage to the reactor core and a failure of the reactor vessel has been estimated at 1 in 20,000 years of reactor operation (i.e., with a fleet of 100 reactors that would be once in 200 years) by the Reactor Safety Study (Rasmussen 1975), and has also been estimated at 1 in 2,000 years of reactor operation by the anti-nuclear Union of Concerned Scientists (Scientists 1977) **[good presentation of opposing normative determinations]**. With well over 2,700 years of commercial reactor operation and over 4,000 years of U.S. Naval reactor

[64] The points in this list do not eliminate particular instances where an individual has some vested interest in nuclear power, which biases their opinions. The point emphasized is that it is quite unreasonable to expect the authoritative reviews and analysis of NAS committees to be systematically biased in a way that contradicts objective scientific results, whether the topic is radiation effects or any other serious matter **[remember the observations of Freudenburg (2001) on how arbitrary decisions or value judgments may get passed off as scientific judgments]**.

operation, there has not been one meltdown accident[65] let alone the 3 expected using the Union of Concerned Scientists calculation. The Reactor Safety Study indicated ~33.5% chance that one would have already occurred.

If a core meltdown accident did occur, there is a distribution of accident severities to be expected. It is reasonable to expect that the containment around a NPP reactor would maintain its integrity for a long time following such an accident, thus preventing any high-radiation exposures or fatalities outside the facility. In 1 out of 5 meltdowns there could be over 1,000 fatalities; in 1 out of 100 there could be over 10,000 fatalities; and in 1 out of 100,000 meltdowns there would be nearly 50,000 fatalities (i.e., on par with annual motor vehicle accident fatalities). Averaging across these types of accident scenarios gives 400 deaths per core meltdown accident by Reactor Safety Study (RSS) estimates and 5,000 deaths by the Union of Concerned Scientists (UCS)[66] **[normative knowledge required here; do I need to review some concepts, especially regarding likelihood estimation?]**. When compared with the estimated 30,000 deaths per year due to air pollution from coal burning, a core meltdown pales in comparison.[67] In fact, it would take 75 meltdowns per year according to the RSS or 6 meltdowns per year by the UCS estimates to equal the current danger posed by coal burning **[Are these data still applicable today, in 2005? Have coal plants become cleaner?]** (Cohen 1990). At this point it might be argued that deaths from air pollution are not dramatic as it is difficult to associate a specific death with coal burning, but the same is true of the death tolls from a meltdown. The vast majority of deaths due to radiation exposure from a meltdown would appear as a slight increase in cancer rates in a large population. Additionally, radiation-induced cancer does not exhibit different physical symptoms than cancer produced by other means. Again Cohen puts this into perspective (Cohen 1990):

> Even in the worst accident considered in the RSS, expected only once in 100,000 meltdowns, the 45,000 cancer deaths could occur among a population of about 10 million, with each individual's risk being increased by 0.5%. Typically, this would increase a person's risk of dying from cancer from 20.0% to 20.5%. This risk varies much more than that from state to state–17.5% in Colorado and New Mexico, 19% in Kentucky, Tennessee, and Texas, 22% in New York, and 24% in Connecticut and Rhode Island–and these variations are rarely, if ever, noticed. It is thus reasonable to assume that the additional cancer risks, even to those involved in this most serious meltdown accident considered in the RSS, would never be noticed **[Was the RSS analysis correct? I must review the assumptions and methodology of this study and more recent studies as well]**.

[65] The Three Mile Island accident which did involve significant melting of the reactor fuel was the closest that the U.S. nuclear power industry has come to a core melt and reactor pressure vessel failure event. Fortunately, analyses performed years after the accident revealed that the pressure vessel never came close to failure (Knief 1992).

[66] The RSS, published in 1975 should be considered a relatively conservative analysis as there have been numerous improvements in plant safety systems and methods of operations that work to further reduce the likelihood of meltdowns (i.e., fuel melt and escape from the primary system) and containment bypasses (i.e., fission products and radionuclides leaving controlled environment of the plant).

[67] It is important to note that this is the estimated number of deaths per year; therefore you have 60,000 deaths after two years, 300,000 deaths after 10 years, 600,000 deaths after 20 years, etc.

In order to directly identify specific deaths attributed to a radiation exposure accident, it is necessary to consider only acute radiation sickness which can be induced by exposures on the order of 500,000 mrems in one day. There were 18 fatalities due to acute radiation sickness world wide[68] (not including the Hiroshima and Nagasaki bombings) up through 1958 largely as a result of the Manhattan project and other weapon related programs (Voelz 1975). Fortunately, there have not been any such fatalities related to acute radiation sickness in the U.S. since that time (Cohen 1990). As mentioned earlier in this paper, there were 31 deaths directly attributed to the Chernobyl accident–many of these involved cases of acute radiation sickness. According to the RSS, there would be no detectable fatalities in 98 of 100 meltdowns, over 100 fatalities in one out of 500 meltdowns, over 1,000 fatalities in one out of 5,000 meltdowns, and in one out of 100,000 core meltdowns there would be roughly 3,500 detectable fatalities (Rasmussen 1975; Cohen 1990).

The record of fatalities for energy related accidents besides nuclear is already at or above many of the levels noted for the very low probability nuclear plant meltdown scenarios listed above. The most dramatic case of detectable fatalities associated with an energy related incident occurred in London in 1952 and was due to air pollution from coal burning. There were 3,500 deaths that occurred within a few days due to this event (Wilson, Colome et al. 1980; Cohen 1990). Figure 7.6 shows fatalities related to energy production by the different energy types from 1979–2004; and Figure 7.7 shows the same data by year. The data for Figures 7.6–7.7 are from the online update to Hore-Lacy (2003) and are conservative for non-nuclear energy sources; as many fatalities go unreported or are very difficult to obtain (e.g., coal mining deaths in China).

[68] The one exception to these worldwide figures has to do with fatalities that may have occurred in the weapons and nuclear power reactor programs of the Former Soviet Union.

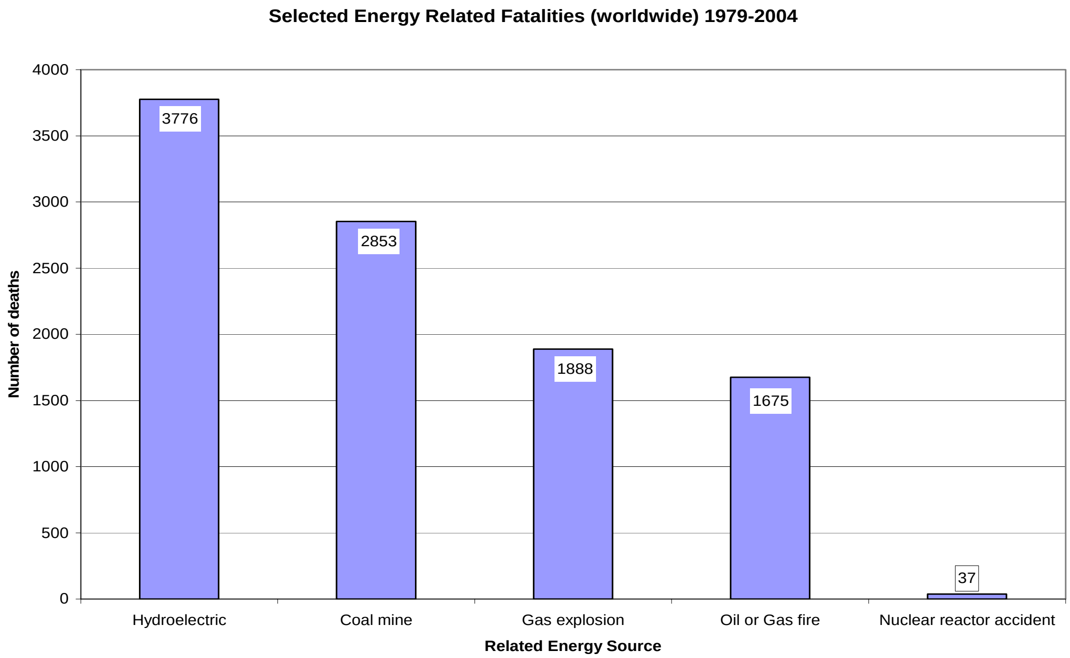

Figure 7.6. Selected energy related fatalities (worldwide) that are directly attributed to energy production. All nuclear energy related fatalities are listed except for ten child fatalities due to thyroid cancer years after the Chernobyl accident in 1986 (Hore-Lacy 2003).

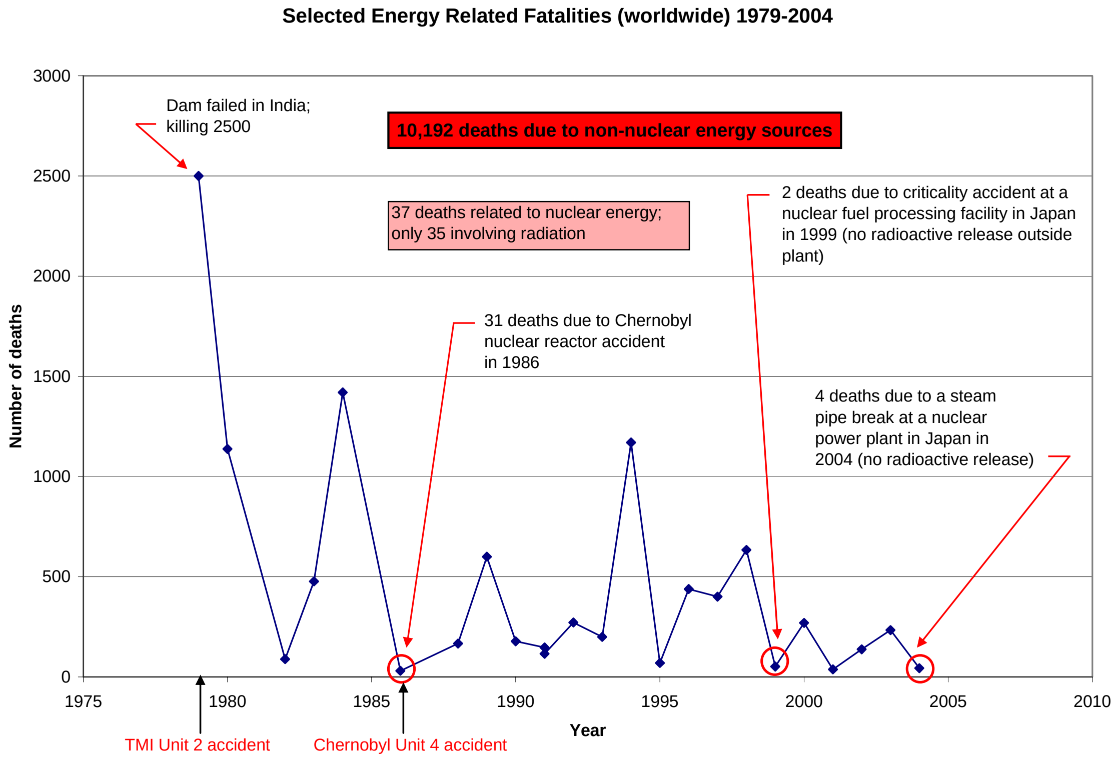

Figure 7.7. Selected energy related fatalities (worldwide) by year from 1979–2004 that are directly attributed to energy production. All nuclear energy related fatalities are listed (Hore-Lacy 2003).

The Uranium Information Center has summarized data from a comprehensive 1998 report completed by the Paul Scherrer Institute in Switzerland that reviewed 4290 energy-related accidents involving over 28,000 fatalities from accidents in different energy sectors (UIC 2004). Figure 7.8 shows deaths per Terawatt-year **[What is a Terawatt?]** for coal (342/ TWe.yr), hydro (883/TWe.yr), gas (85/TWe.yr), and nuclear power (8/TWe.yr). The data for Figure 7.8 covers accidents during the period from 1969–1996.

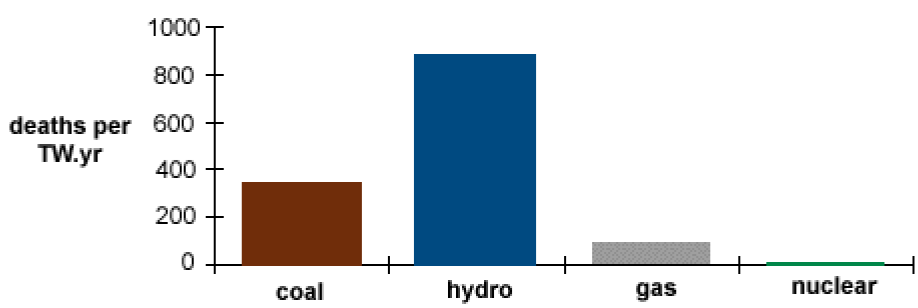

Figure 7.8. Deaths per TWe.yr in different energy production sectors (UIC 2004) **[This is an interesting figure. Although, with Hydro, I wonder how many deaths have been prevented by hydroelectric plants which have, by and large, reduced the cyclical flooding of many populated areas?]**.

The existing data regarding fatalities and energy production definitely reveal nuclear power at the safest electric power generation industry to date. But one factor makes nuclear power accidents and resulting injury risks a bit different than for the other sources–this factor is persistence of danger following a major accident. That is, large areas of land may end up contaminated with harmful fission decay products for many years following an accident. Figure 7.9 below shows a Cesium-137 (Cs-137) contamination map and related access control zones around the Chernobyl accident site produced by the U.S. Central Intelligence Agency. The spatial separation of highly contaminated areas was due to lofting of Cs-137 high into the air by the explosion, and then subsequent rainstorms washed the particles to the ground; where it rained is where the concentrations are found. The half life of Cs-137 is 30.2 years and it goes through beta decay, which produces Barium-137, which then decays and releases a gamma ray. Other energy sources do not carry with them quite this type of persistence of danger **[What about the persistence of mercury, cyanide and other toxins? Many pollutants do not have a half-life, they are always a hazard.]**. There are major environmental impacts of oil spills which contaminate water and destroy animal populations, and pollution effects from fossil fuel burning include shortened lives and increased incidence of respiratory ailments. But the type of persistent damage caused by a release of radioactive fission products seems especially unnerving for many people. This is likely due to phenomenological familiarity issues as described in section four of this report **[the general public may have significant gaps in substantive knowledge on this topic; perceived nature of death & dread effects are important here; I seem to recall the recent television program that discussed the types of radioactive materials present in fly ash from fossil fuel power plants–I need to learn more about that issue as well]**.

It is interesting to note, however, that certain long-time environmentalists or 'green' activists have recently declared that greatly expanding nuclear power is the only rational approach to combat environmental and health impacts of other energy sources. James Lovelock, a climatologist and self-proclaimed 'Green' who sees the dangers from global warming to be imminent has argued forcefully that the world should focus on nuclear power as a safe, environmentally friendly way to supply the vast majority of electricity for the developed world. Although he does have hope for improved efficiencies and reduced costs for wind, solar, and hydroelectric in the future, he sees no other reasonable alternative than to transition rapidly to nuclear power in order to reduce greenhouse gas emissions that are exacerbating global warming[69] (Lovelock 2004). Climatologist Stewart Brand, founder of the Whole Earth Catalog, and the co-founder of Greenpeace, Patrick Moore, have also endorsed nuclear power as a credible, clean alternative to natural gas and coal fired power plants (Thompson 2005).

[69] It appears that James Lovelock's recent strong support for nuclear power to reduce the human induced global warming was triggered by the tremendous heat wave in Europe during 2003 that has been blamed for as many as 35,000 deaths. Lovelock referred to this meteorological event several times in his article.

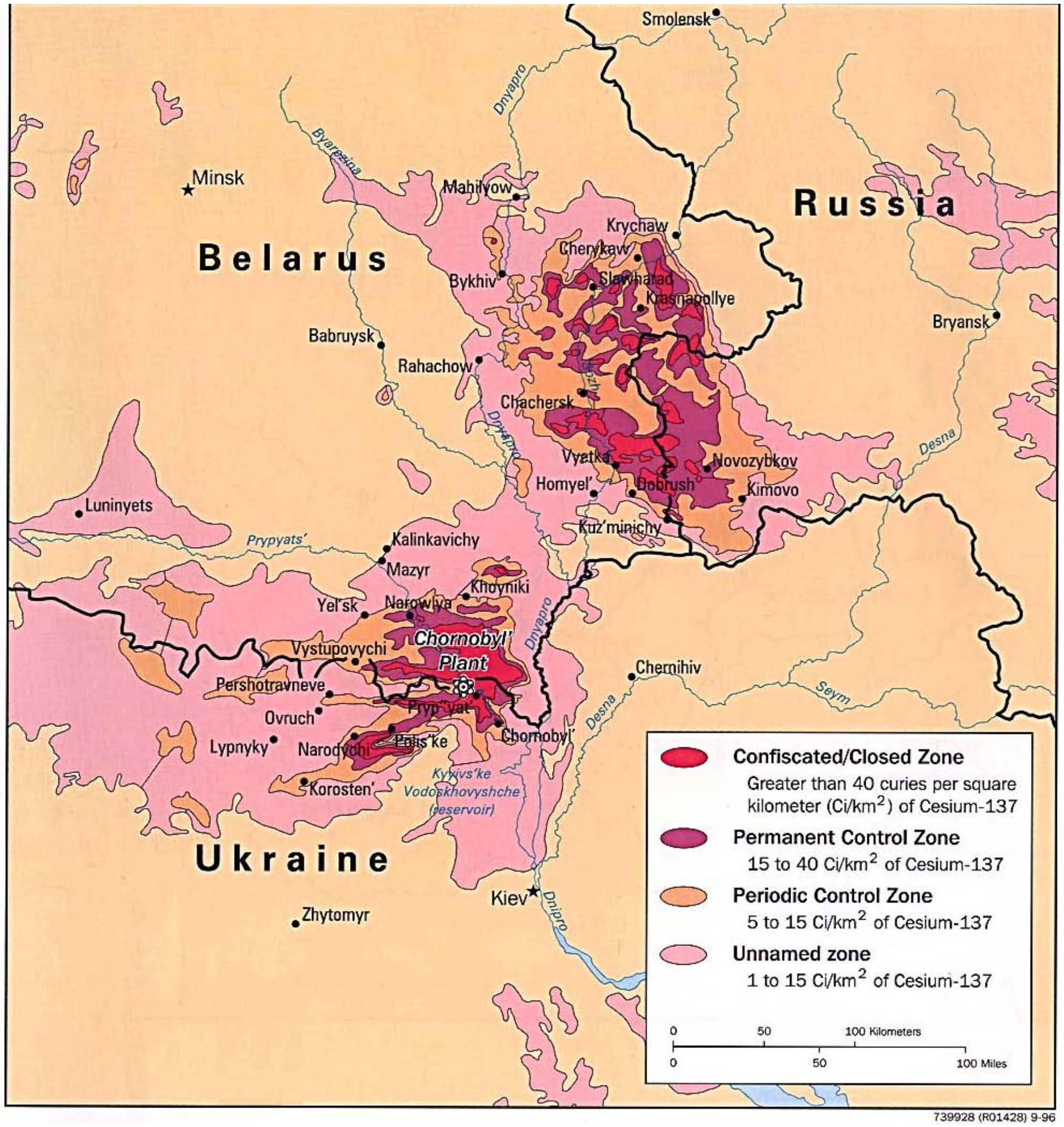

Figure 7.9. Distribution of persistent Cs-137 contamination at and surrounding the Chernobyl Unit 4 explosion. Image originally created by the U.S. Central Intelligence Agency and published in the CIA Handbook of International Economic Statistics (1996) and made available through the Perry-Casteñeda Library Map Collection at The University of Texas at Austin, http://www.lib.utexas.edu/maps/belarus.html.

At this point, notwithstanding the small quantitative 'risks' presented regarding radiation releases from nuclear power plants and the very impressive overall record of safety for light water reactors to date, the reader may recall the earlier sections of this report in which known unknowns, unknown unknowns, aleatory and epistemic uncertainty, and specific references to potential problems with quantitative risk assessment models were mentioned. Such recollections may lead to the very logical question: "How does the U.S. nuclear power industry defend against inevitable uncertainties in their systems analysis models to ensure protection of people and the environment from radioactive releases?"

The answer to this question lies in the defense-in-depth (DID) approach which has been informally adopted by the U.S. nuclear power industry and by many in the United States Nuclear Regulatory Commission (USNRC).

Defense-in-depth refers to a three-level approach consisting of *prevention*, *protection*, and *mitigation*. Prevention seeks to completely avoid the operational occurrences that could result in damage, loss of fuel performance, abnormal release of radioactivity, or other accident precursors. This requires high reliability components, systems, and operating practices. Further measures include inherent stability of the nuclear fission process in the system design, safety margins in operations, testing and inspection, regular training, quality assurance to ensure adherence to policies and procedures, etc. (Knief 1992). One of the simplest definitions of DID is the following: “Multiple barriers of increasing conservatism” (Powers 2005).

In addition to prevention, the nuclear power industry and the NRC have recognized that some component failures and operating errors will be inevitable over the lifetime of a nuclear reactor (or, any other technologically advanced system involving people and equipment). Therefore, a line of defense called *protection* is employed. This involves measures that halt or otherwise handle low-probability incidents and transients that cause a shutdown of the fission process and possibly to minor fuel damage and small radioactive releases. Many conceivable failures are postulated and analyzed so that reliable protective measures may be employed. Examples of protective measures in a nuclear power reactor plant include: use of control rods to quickly halt the fission reaction; pressure relief valves; interlocks; automatic monitoring and safety-system initiation; transient operating guidelines and procedures; continuous monitoring and control of radiation levels; etc (Knief 1992).

The third level of DID consists of *mitigation* systems. If a very rare core damage accident were to occur, these systems act to limit the consequences to people and the environment. Examples of mitigation systems include: emergency water supplies to substitute from primary water supplies; high and low pressure water systems specifically targeting the reactor core; back-up electrical power supplies to power an array of critical systems; massive steel or steel reinforced concrete[70] containment structures; and emergency planning and procedures (Knief 1992). Figure 7.10 provides a summary of the DID approach use in the U.S. nuclear power industry.

[70] The type of steel reinforced concrete construction for nuclear power plant containment structures is extremely unique. First, a mesh of steel members is assembled that is so tightly woven, it nearly prevents all direct light from passing. Second, specially designed pressurized concrete pump systems must force the concrete in and around the dense steel framework (Cohen 1990) **[are all reactor containments built this way? Must check]**.

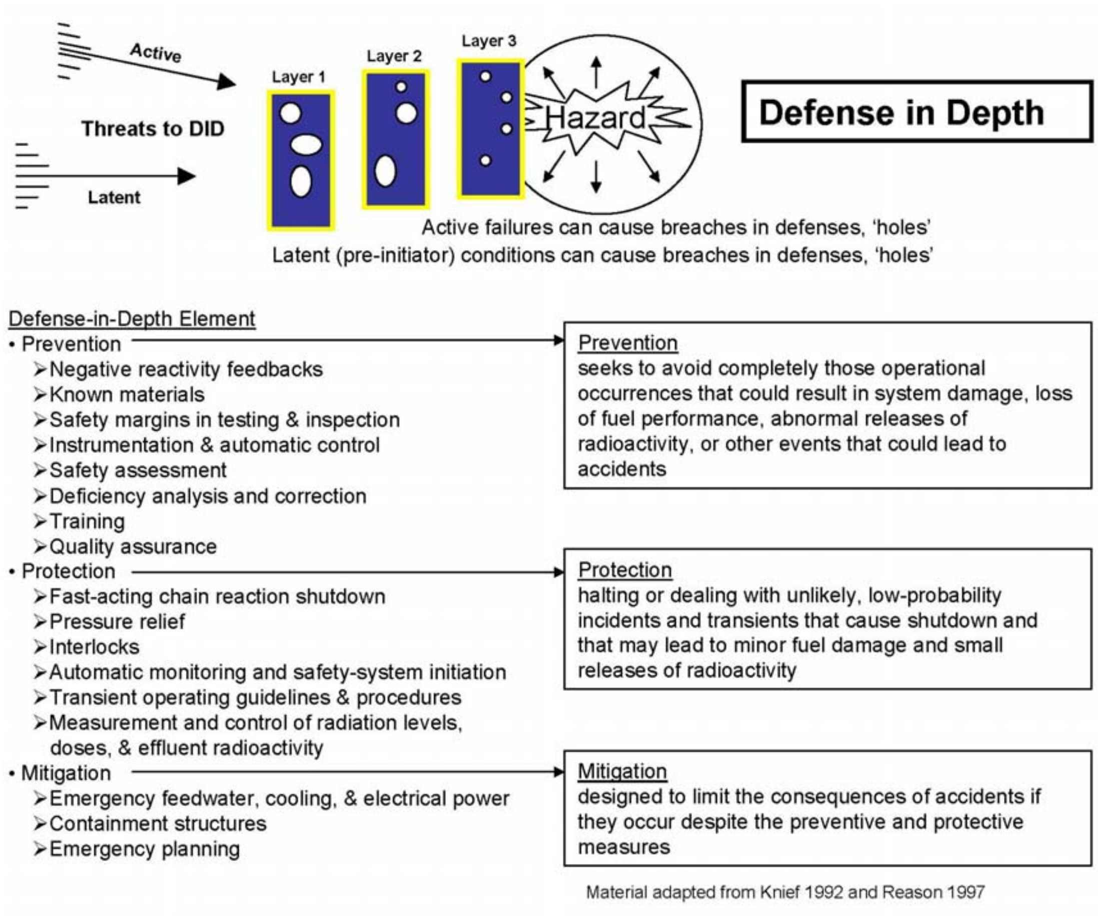

Figure 7.10. U.S. defense-in-depth approach for protecting people and the environment from radioactive releases from nuclear power plants. Adapted from Knief (1992) and Reason (1997).

Now that a body of information has been presented in support of the claim that the risks from generating electricity with nuclear power are much lower than for several alternatives, let us consider the immensely important decision of whether or not to expand the role of nuclear power in electricity generation. To do this, it is critical to gauge the opinions and sentiments of those affected by such decision. In this case, the general U.S. public would be among those affected **[members of the public should be directly involved in the decision making process]**. This means that an understanding of current perceptions among the U.S. public regarding nuclear power is essential.

According to one recent survey, the general public in the United States may not be connecting concern about global warming with carbon-free nuclear power. An internet-based survey of 1,350 adults in the U.S. revealed the following insights with respect to nuclear power (Ansolabehere, Deutch et al. 2003):

1. On average respondents preferred that the U.S reduce somewhat nuclear power, coal, and oil while keeping natural gas at its current level
2. Respondents strongly support a significant expansion of wind and solar power

3. Perceived environmental harm due to nuclear power was the most important factor for wanting a decrease in nuclear power
4. Safety and waste disposal issues comprised the second most important group of factors for wanting a decrease in nuclear power
5. Perceived costs were the third most important factor for wanting a decrease in nuclear power
6. Concern about global warming does not predict preferences about future use of nuclear power
7. Political beliefs, and demographics including age, gender, and income mattered very little with regard to responses

Interestingly, when respondents were shown the relative costs for non-carbon emitting sources of power (i.e., wind and solar were shown to be more expensive than nuclear); many chose nuclear power as the preferred option. Therefore, educating the public with regard to costs **[Do the typical cost estimates provided by the nuclear industry include the large up-front capital costs?]** may increase support for the expansion of nuclear power. It should be noted that the methodology used in the internet-based MIT survey likely does not provide a representative sample of electricity consumers, *so these results should be interpreted with caution.*

Late in 2004, the Institute of Electronic and Electrical Engineers (IEEE) published an online article discussing current attitudes regarding nuclear power, and they provided accumulated survey data over a number of years into several helpful graphs. Figure 7.11 shows the percentage of Americans who have supported, opposed, or don't know whether the United States should build new nuclear power plants in surveys conducted over the past 25 years. Asking the public whether they favor building new plants appears to be a reliable measure for gauging the overall acceptance of nuclear power according to the researchers interviewed in the IEEE article (Miller 2004). Figure 7.12 reveals the differences that appear in survey results depending upon the specific questions asked. For example, if the question is asked in this way, "Which of the following is the best solution to U.S. energy problems?" then nuclear power does not fare well. If an accident related question appears, e.g., "Do you think a situation such as the nuclear power plant breakdown at Three Mile Island is likely to occur in the future?" then the response is typically in the affirmative.[71] When questions about a specific energy crisis are tied to nuclear power alternative, then nuclear power tends to fare quite well in surveys (Miller 2004) **[this really helps to understand the specific framing effects that are important here; framing in different ways helps people to consider 'realized bad things' as opposed to just 'potential bad things' as being more worrisome; availability biases are definitely at work here]**.

[71] It should be mentioned that when a question such as this is asked, there is no time limit to the time period implied. Therefore, it is very sensible to think that a major plant breakdown will occur at some point in the future.

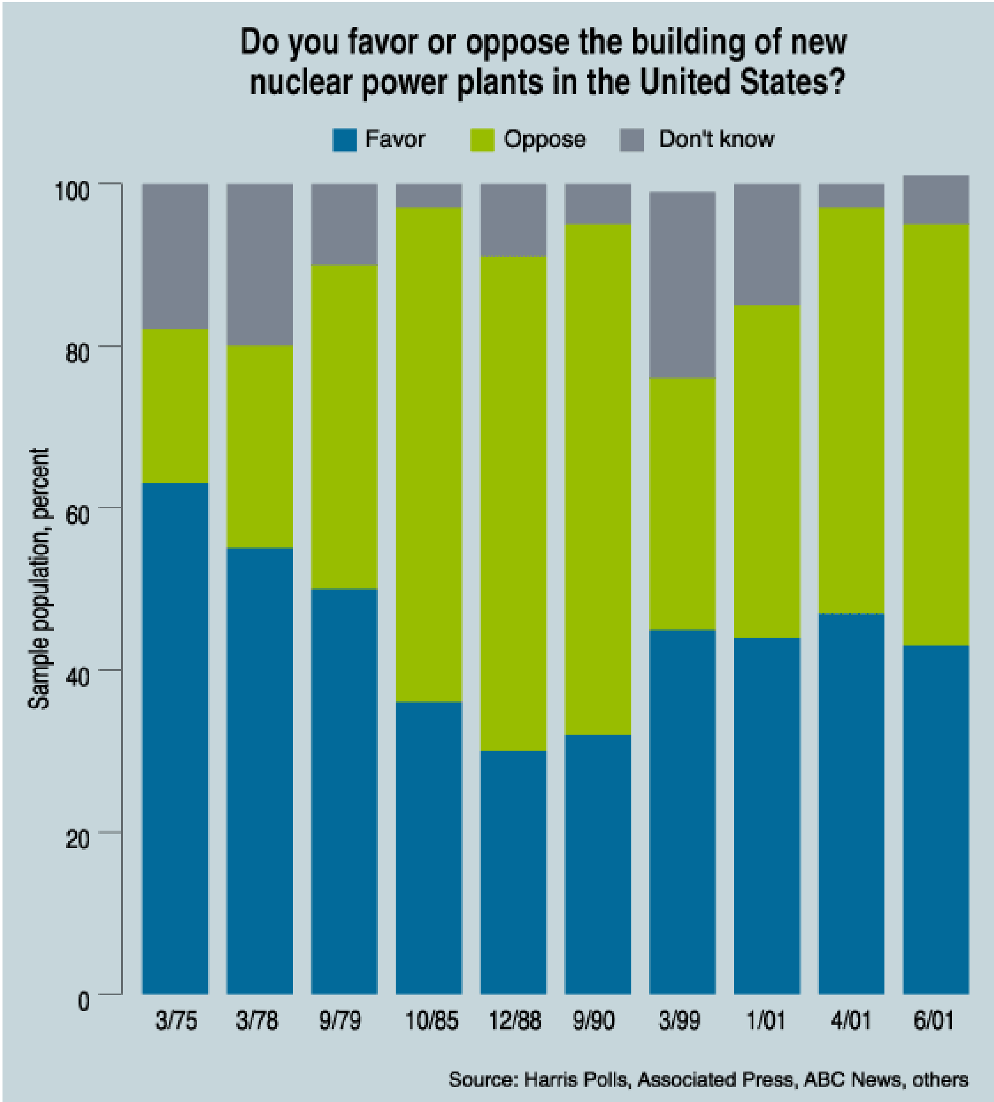

Figure 7.11. Percentage of Americans who favor, oppose, or don’t know whether the United States should build new nuclear power plants (Miller 2004).

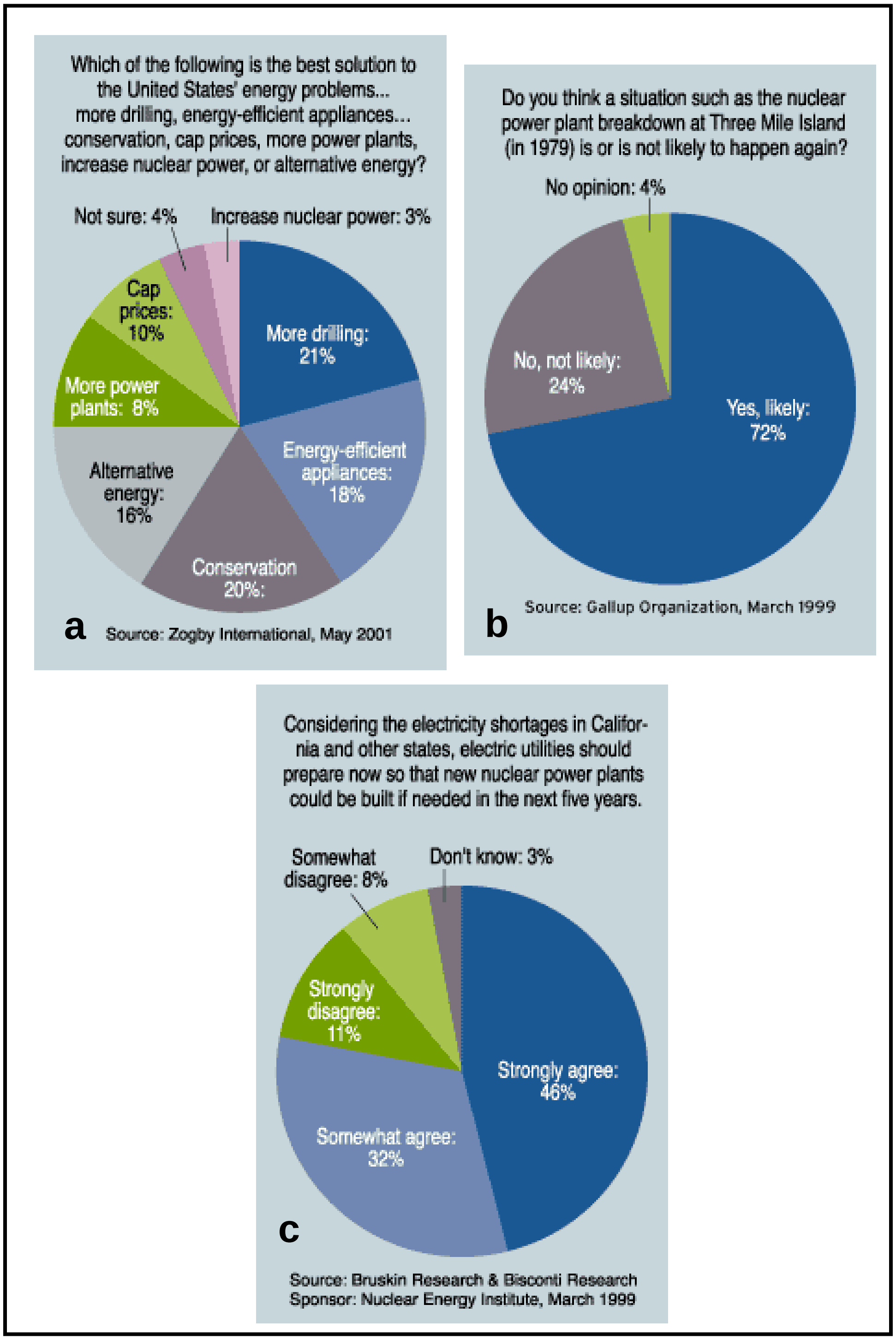

Figure 7.12. Alterations in the percentage of people supporting nuclear power as a function of question presentation (Miller 2004).

It is hoped that the data and information provided in this section have achieved the stated goals of demonstrating that the risks from generating electricity with nuclear power are much lower than for several alternatives, and showing that previous heightened public fears regarding nuclear power are principally due to biased media coverage. Furthermore, data have been provided showing that support for building new nuclear power reactors is growing. Currently, a majority of people in the U.S. see the expansion of nuclear power as a viable option for ensuring an economical, abundant **[Over what time scales will there be abundant supplies? Must research this.]**, safe, and reliable supply of electricity **[The real goal for this section has been to provide the reader with an example opportunity to see how decision making bias awareness and bias mitigation strategies may be applied to a specific source of data used as a technical input to the decision making process regarding high-consequence, technologically advanced decisions]**.

8. Summary of Unique Framework and Decision-Making Approach

As stated at the beginning of this report, it is suggested that all decision makers, whether ordinary citizens, academics, or political leaders, ought to cultivate their abilities to participate in decisions regarding high-consequence, highly sophisticated technological systems. Key to this process is understanding the factors that influence risk perception and strategic decision making. This report has brought together insights regarding risk perception and decision making across domains ranging from nuclear power technology safety, cognitive psychology, economics, science education, public policy, and neural science (to name a few) and form them into a unique, coherent, concise framework and list of strategies to aid in decision making. Many of the specific examples provided are tailored to the issue of using nuclear power to generate electricity. Fortunately, the decision making framework and approach presented here are applicable to any high-consequence, highly sophisticated technological system. The goal of this report is to promote increased understanding of decision making processes and hopefully to enable improved decision making.

In covering this complex topic, the following items were presented:

- A unique framework for understanding decision making was briefly presented along with a couple of illustrative examples
- Critical thinking processes considered foundational to proper decision making were defined and described
- The nature of a ‘technologically advanced’ society (relevant to decision making) was briefly discussed
- A detailed discussion (including examples) of the foundations for the framework was presented
- A step-by-step description of a recommended decision making approach was presented
- A discussion of perceived risks and analyzed risks related to nuclear power was presented along with pointers to the proposed framework and decision making approach

Finally, the all important question is restated: “If the decision maker or decision making team follows the process presented here, will it lead to better decisions?” The answer supported here is that following the process will lead to better decisions in nearly all cases, and the outcomes of those decisions will typically be superior to decisions made without the structured, bias-aware approach. That said, remember these two points: first, good decisions sometimes lead to bad outcomes and bad decisions sometimes lead to good outcomes; second, bad outcomes tend to be remembered much more easily than good outcomes–they are more *available*. Thus the prudent, disciplined decision maker(s) must stay the course and tally the results over time to realize the fruit of the highly-structured decision making approach when decisions regarding high-consequence, highly sophisticated technological systems are required. As a parting note, please reflect on the re-presentation of the approach to high-consequence, technologically advanced decisions proposed in this report (see Figure 8.1 below).

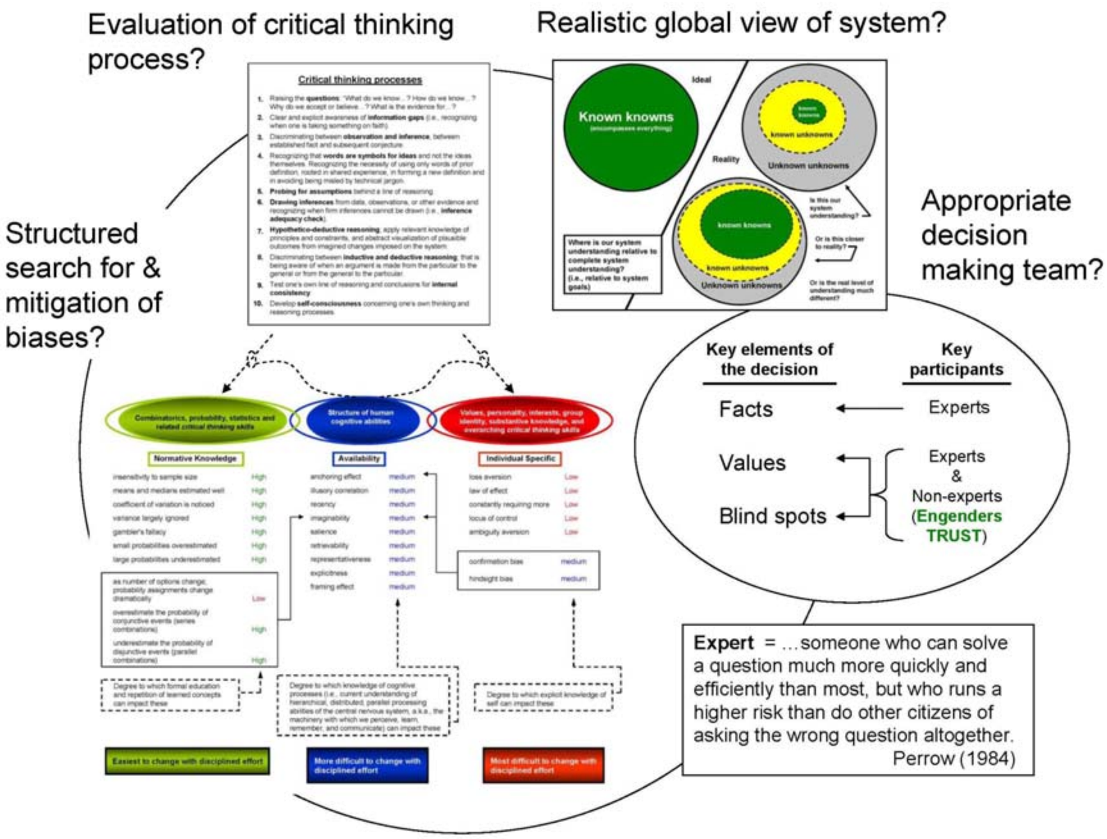

Figure 8.1. Overview of the approach to high-consequence, technologically advanced decisions.

9. Appendix

Given any two events *A* and *B*, such as, the event of a NPP meltdown (*A*), and the event of a loss of coolant accident in a NPP (*B*), recall that:

$$p(A \mid B) = \frac{p(A \cap B)}{p(B)} \Rightarrow p(A \cap B) = p(A \mid B)p(B) = p(B \mid A)p(A), \tag{1}$$

$$\therefore p(A \mid B) = \frac{p(B \mid A)p(A)}{p(B)}, \tag{2}$$

substituting *H* for *A* and *D* for *B* – where *H* represents hypothesis, underlying truth or cause and *D* represents data observed, apparent symptoms, or effect – results in:

$$p(H \mid D) = \frac{p(D \mid H)p(H)}{p(D)}. \tag{3}$$

This means that once some data are available for a case of known *H*, one is able to refine any prior estimate of probability of H, *p*(*H*), by multiplying by *p*(*D*|*H*) and dividing by *p*(*D*). The term *p*(*D*|*H*) is what is known from prior experience, (i.e., where one knew *H* to be true and then observed *D*) and it is often calculated using a statistical model. The term *p*(*D*) is the probability of those data being observed in general, independent of other circumstances. Now it is important to point out that obtaining the values of *p*(*D*|*H*) for each *D* and *H* is the critical step in applying Bayes's theorem to any revision of opinion in light of information. Also, the hypothesis *H* is one of a list, or partition, of mutually exclusive and exhaustive hypotheses H_i. Since $P(D \cap H_i) = P(D \mid H_i)P(H_i)$ and since

$$\sum_i P(D \cap H_i) = P(D), \tag{4}$$

we then have

$$P(D) = \sum_i P(D \mid H_i)P(H_i).$$

It may be very difficult to acquire data for P(D), but another way of structuring the problem can relieve this difficulty. Suppose you observe D_1 and use Equation (3) to refine the prior knowledge *p*(*H*) to get $p(H \mid D_1) = [p(D_1 \mid H)p(H)/p(D_1)]$. Then you observe D_2. It is now possible to use the equation again to determine the new knowledge of *H*, $p(H \mid D_1 D_2)$, by substituting $p(H \mid D_1)$ as the current prior knowledge of *H*,

$$p(H \mid D_1 D_2) = \frac{p(D_2 \mid H)\left[\frac{p(D_1 \mid H)p(H)}{p(D_1)}\right]}{p(D_2)}, \tag{4}$$

where the term in the brackets is $p(H \mid D_1)$.

Now suppose from the beginning you have two independent hypotheses, H_1 and H_2, and following the process shown in Equation 4. You would now be able to take a ratio of the two equations in which $p(D_1)$ and $p(D_2)$ would fall out since they are common to both equations:

$$\frac{p(H_1 \mid D_1 D_2)}{p(H_2 \mid D_1 D_2)} = \frac{p(D_2 \mid H_1)}{p(D_2 \mid H_2)} \cdot \frac{p(D_1 \mid H_1)}{p(D_1 \mid H_2)} \cdot \frac{p(H_1)}{p(H_2)}. \qquad (5)$$

The first term in Equation 5 is referred to as the *posterior odds ratio*. The last term is called the *prior odds ratio*. The two terms in between are referred to as the *likelihood ratios* for D_2 and D_1, respectively. This process can be extended to as many data as one wishes in order to find a posterior odds ratio. It is important to note that the data do not have to be independent of one another. The general result is that the posterior odds ratio is the product of likelihood ratios for all salient data multiplied by the prior odds ratio. This approach is independent of the order in which the data are observed. It does make the assumption that the process statistics are stationary (i.e., they don't change over time), which is a powerful assumption that may not always be true for a given system.[72] Old data and new data can be used; neither is better than the other if the stationary process assumption is valid. If the process is found or believed to be slightly non-stationary, the updating process can be adjusted to discount data according to its age. Therefore, the Bayesian approach to updating event probabilities is a logically consistent technique that can be very effective for using data to reduce uncertainty.

A simple example of comparing the Bayesian decision approach to that of a 'typical' person includes the following: Let (C) be the hypothesis that a NPP core meltdown occurs in a given year, let (R_1) be the failure of a critical reactor component, and let (R_2) equal the failure of a source of electrical power to the NPP. Given the following probabilities:

- The probability that a NPP experienced a failure of the critical reactor component given that a core meltdown occurred soon after that reactor component failed (i.e., certainly within the year and there is good reason to suspect a causal link) is 0.4 (i.e., $p(R_1|C) = 0.4$)
- The probability that a NPP experienced a failure of the critical reactor component during the year given that a core meltdown did not occur is 0.005 (i.e., $p(R_1|C^c) = 0.005$).
- The probability that a NPP experienced a failure of the electrical power source given that a core meltdown occurred soon after the electric power failed is 0.8 (i.e., $p(R_2|C) = 0.8$)
- The probability that a NPP experienced a failure of the electrical power source during the year given that a core meltdown did not occur is 0.4 (i.e., $p(R_2|C^c) = 0.4$).
- The probability that a core meltdown occurs is 0.00005 (i.e., $p(C) = 0.00005 \rightarrow p(C^c) = 0.99995$).

[72] Of course, classical statistical approaches typically assume independent, identically distributed random variables when inferring population parameters from samples.

What is the probability that a core meltdown occurs given that a critical reactor component fails (i.e., $p(C|R_1)$)?

$$p(C \mid R_1) = \frac{p(R_1 \mid C)p(C)}{p(R_1)} = \frac{p(R_1 \mid C)p(C)}{[p(R_1 \mid C)p(C)] + [p(R_1 \mid C^c)p(C^c)]} ,$$

$$\Rightarrow \frac{(0.4)(0.00005)}{[(0.4)(0.00005)] + [(0.005)(0.99995)]} = 0.003984 .$$

That is, if this situation happened 1,000 times, 4 core meltdowns would be expected. If a person were simply given these probabilities and asked to determine the probability of a core meltdown given that the critical reactor component had failed, the answer would likely be much closer to 40% than to 0.4%.

Continuing further, what would be the probability of having a core meltdown given that both the critical reactor component failed and the electrical power was lost?

$$p(C \mid R_1 R_2) = \frac{p(R_2 \mid C)\left[\dfrac{p(R_1 \mid C)p(C)}{p(R_1)}\right]}{p(R_2)} = \frac{p(R_2 \mid C)\left[\dfrac{p(R_1 \mid C)p(C)}{[p(R_1 \mid C)p(C)] + [p(R_1 \mid C^c)p(C^c)]}\right]}{[p(R_2 \mid C)p(C)] + [p(R_2 \mid C^c)p(C^c)]},$$

$$\Rightarrow \frac{(0.8)(0.003984)}{[(0.8)(0.00005)] + [(0.4)(0.99995)]} = 0.007968 .$$

This result tells us that having the additional knowledge that the power went out on top of the failure of the critical component only increased the likelihood that a core meltdown would occur by roughly 0.4%. Based on the human tendencies mentioned previously, it is likely that a person would take the $p(R_1|C) = 0.4$ and the $p(R_2|C) = 0.8$ and arrive at a value for core damage that is much higher than approximately 0.8%. After going through the machinations of the Bayesian example above, it may be apparent why such an approach is not readily used in the course of human decision making. It requires both specific training and experience to apply and generally requires a disciplined accumulation of data. Given that data do not support the hypothesis that an innate Bayesian estimation tendency exists in people, one must exert considerable effort in mastering the technique. Yet, it is argued here that Bayesian techniques are a very helpful way in which to make maximum use of the data that are available for making a decision and should be the standard against which people's inferential tendencies may be measured (Sheridan and Ferrell 1974). The interested reader is encouraged to consult Winterfeldt and Edwards (1986) for an accessible treatment of Bayesian statistics versus classical statistics.[73]

---

[73] The author of this paper supports the use of Bayes' Theorem and its extensions as an appropriate method for quantifying well-structured inferences. See Winterfeldt and Edwards (1986) and Sheridan and Ferrell (1974).

10. References

Airsafe.com, L. (2004). Fatal events and fatal event rates by airline since 1970. http://www.airsafe.com/airline.htm.

Anderson, J. R. (1980). Cognitive Psychology and Its Implications. San Francisco, Freeman.

Anderson, N. H. and A. Jacobson (1965). "Effect of stimulus inconsistency and discounting instructions in personality impression formation." Journal of Personality and Social Psychology **2**: 531-539.

Ansolabehere, S., J. Deutch, et al. (2003). The future of nuclear power; an interdisciplinary MIT Study. Cambridge, MA, Massachusetts Institute of Technology.

Arons, A. B. (1990). A Guide to Introductory Physics Teaching. New York, John Wiley & Sons, Inc.

Bernoulli, D. (1738). "Specimen Theoriae Novae de Mensura Sortis (Exposition of a New Theory on the Measurement of Risk)." Econometrica **22**(1954): 23-36.

Bernstein, P. L. (1998). Against the Gods: The Remarkable Story of Risk. New York, NY, John Wiley & Sons.

Bower, T. G. R. (1966). The visual world of infants. Readings from Scientific American: Perception: Mechanisms and Models (collection of articles published as book in 1971). San Francisco, W.H. Freeman and Company**:** 349-357.

Brady, J. E. and J. R. Holum (1984). Fundamentals of Chemistry. New York, John Wiley & Sons.

Budziszewski, J. (1999). The Revenge of Conscience: Politics and the Fall of Man. Dallas, TX, Spence Publishing Company.

Cha, Y. J. (2004). "An analysis of nuclear risk perception: with focus on developing effective policy alternatives." International Review of Public Administration **8**(2): 33-47.

Cha, Y. J. (2000). "Risk perception in Korea: a comparison with Japan and the United States." Journal of Risk Research **3**(4): 322-332.

Clark, L. and J. F. S. Jr. (1993). "Social organization and risk: some current controversies." Ann. Rev. Sociology **19**: 375-399.

Cohen, B. L. (1990). The Nuclear Energy Option. New York, NY, Plenum Press.

Cooper, A. and R. Robinett III (2005). "Analyzing and Overcoming Fear in Personnel Performance and Interpersonal Interactions." Journal of System Safety **(submitted for publication)**.

Conover, W. J. (1999). Practical Nonparametric Statistics. New York, NY, John Wiley & Sons, Inc.

Covan, J. and J. A. Cooper (2004). A tutorial on top-down system safety. 22nd Annual International System Safety Conference. Providence, RI, The Systems Safety Society.

Edwards, W. (1968). Conservatism in human information processing. Formal Representation of Human Judgment. B. Kleinmutz. New York, NY, John Wiley & Sons, Inc.**:** 17-52.

Einhorn, H. J. and R. M. Hogarth (1978). "Behavioral decision theory: Processes of judgment and choice." Annual Review of Psychology **32**: 53-88.

Erakar, S. E. and H. C. Sox (1981). "Assessment of patients' preferences for therapeutic outcomes." Medical Decision Making **1**: 29-39.
Fantz, R. L. (1961). The origin of form perception. Readings from Scientific American: Perception: Mechanisms and Models (collection of articles published as book in 1971). San Francisco, W.H. Freeman and Company**:** 334-340.
Fink, H. (1981). Social Philosophy. New York, Methuen.
Fischhoff, B. (1983). "Predicting frames." Journal of Experimental Psychology: Learning, Memory and Cognition **9**: 103-116.
Fischhoff, B., P. Slovic, et al. (1980). Knowing what you want: Measuring labile values. Cognitive processes in choice and decision behavior. T. Wallsten. Hillsdale, NJ: Erlbaum**:** 117-141.
Fischoff, B. (1975). "Hindsight = foresight: The effect of outcome knowledge on judgment under uncertainty." Journal of Experimental Psychology: Human Perception and Performance **1**: 288-299.
Fischoff, B., P. Slovic, et al. (1978). "How safe is safe enough? A psychometric study of attitudes toward technological risks and benefits." Policy Science **9**: 127-152.
Fishburn, P. C. and G. A. Kochenberger (1979). "Two-piece von Newmann-Morgenstern utility functions." Decision Sciences **10**: 503-518.
Fox, C. R. and A. Tversky (1995). "Ambiguity aversion and comparative ignorance." Quarterly Journal of Economics **CX**(3): 585-603.
Freudenburg, W. R. (1993). "Weber, the division of labor, and the rationality of risk perceptions." Social Forces **71**(4): 909-932.
Freudenburg, W. R. (2001). "Risky thinking: facts, values and blind spots in societal decisions about risks." Reliability Engineering and System Safety **72**: 125-130.
Geertz, C. (1973). The Interpretation of Cultures. New York, Basic Books.
Gibson, E. J. and R. D. Walk (1960). The visual cliff. Readings from Scientific American: Perception: Mechanisms and Models (collection of articles published as book in 1971). San Francisco, W.H. Freeman and Company**:** 341-348.
Glasstone, S. and A. Sesonske (1963). Nuclear Reactor Engineering. Princeton, D. Van Nostrand Company, Inc.
Greenwald, A. G., A. R. Pratkanis, et al. (1986). "Under what conditions does theory obstruct research progress?" Psychological Review **93**: 216-229.
Hacking, I. (1975). The Emergence of Probability: A Philosophical Study of Early Ideas About Probability, Induction, and Statistical Inference. London, Cambridge University Press.
Halliday, D. and R. Resnick (1988). Fundamentals of Physics: Third Edition Extended. New York, John Wiley & Sons.
Hershey, J. C. and P. J. H. Schoemaker (1980). "Risk taking and problem context in the domain of losses: An expected-utility analysis." Journal of Risk and Insurance **47**: 111-132.
Heuer, R. J. (1982). Cognitive factors in deception and counterdeception. Strategic Military Deception. D. C. Daniel and K. L. Herbig. New York, Pergamon Press**:** 31-69.
Hillel, M. B. (1973). "On the subjective probability of compound events." Organizational Behavior and Human Performance **9**: 396-406.
Hines, W. W. and D. C. Montgomery (1990). Probability and Statistics in Engineering and Management Science. New York, NY, John Wiley & Sons.

Hogarth, R. M. (1975). "Cognitive process and the assessment of subjective probability distributions." Journal of the American Statistical Association **70**(350): 271-289.
Hora, S. C. (1996). "Aleatory and epistemic uncertainty in probability elicitation with an example from hazardous waste management." Reliability Engineering and System Safety **54**: 217-223.
Hore-Lacy, I. (2003). Nuclear Electricity. Melbourne, Australia, Uranium Information Centre Ltd. and World Nuclear Association.
IAEA (1991). IAEA technical report. Vienna, Austria, International Atomic Energy Agency.
Inhaber, H. (2004). Risk analysis applied to energy systems. Encyclopedia of Energy. C. Cleveland, R. U. Ayres, R. Costanzaet al. St. Louis, MO, Elsevier.
Jasanoff, S. (1998). "The political science of risk perception." Reliability Engineering and System Safety **59**: 91-99.
Juran, J. M. (1989). Juran on Leadership for Quality. New York, Free Press.
Juran, J. M. and F. M. Gryna (1988). Quality Control Handbook. New York, McGraw-Hill.
Kahneman, D. and A. Tversky (1979). "Prospect Theory: An analysis of decision under risk." Econometrica **47**(2): 263-291.
Kahneman, D. and A. Tversky (1984). "Choices, Values, and Frames." American Psychologist **39**(4): 341-350.
Kandel, E. R., I. Kupfermann, et al. (2000). Learning and memory. Principles of Neural Science. E. R. Kandel, J. H. Schwartz and T. M. Jessell. New York, McGraw-Hill**:** 1227-1246.
Kandel, E. R., J. H. Schwartz, et al., Eds. (2000). Principles of Neural Science. New York, McGraw-Hill.
Kandel, E. R. and R. H. Wurtz (2000). Constructing the visual image. Principles of Neural Science. E. R. Kandel, J. H. Schwartz and T. M. Jessell. New York, McGraw-Hill**:** 492-506.
Klein, G. (1998). Sources of Power: How People Make Decisions. Cambridge, MA, MIT Press.
Klein, G. and B. W. Crandall (1995). The role of mental simulation in naturalistic decision making. Local applications of the ecological approach to human-machine systems. P. Hancock, J. Flach, J. Caird and K. Vicente. Hillsdale, NJ, Erlbaum.
Knief, R. A. (1992). Nuclear Engineering: Theory and Technology of Commercial Nuclear Power. Washington, DC, Taylor & Francis.
Kolarik, W. J. (1995). Creating Quality: Concepts, Systems, Strategies, and Tools. New York, McGraw-Hill.
Larsen, R. J. and M. L. Marx (2001). An Introduction to Mathematical Statistics and Its Applications. Upper Saddle River, NJ, Prentice Hall.
Miller, G. (1956). "The magical number seven plus or minus two: Some limits on our capacity for processing information." Psychological Review **63**: 81-97.
Miller, G. A. (1962). Psychology: The Science of Mental Life. New York, Harper & Row.
Miller, S. (2004). "Pragmatic Concerns Fuel Nuclear Support." Retrieved November 1, 2004, from www.spectrum.ieee.org.
Mish, F. C., Ed. (1996). Merriam Webster's Collegiate Dictionary. Springfield, MA, Merriam-Webster, Inc.

Mokdad, A.H., J.S. Marks, D.F. Stroup, and J.L. Gerberding (2004). "Actual causes of death in the United States, 2000" The Journal of the American Medical Association **291**(10): 1238-1245.

NCRP - National Council on Radiological Protection (1975). Natural background radiation in the United States.

Neumann, J. and O. Morgenstern (1944). Theory of Games and Economic Behavior. Princeton, NJ, Princeton University Press.

Nolan, D. (1997). "Quantitative Parsimony." British Journal for the Philosophy of Science **48**(3): 329-343.

NUREG/CR-6823 (2003). Handbook of Parameter Estimation for Probabilistic Risk Assessment. Washington, DC, Nuclear Regulatory Commission.

Orwell, G. (1993). Animal Farm. New York, Everyman's Library.

Ozkaynak, H. and J. C. Spengler (1985). "Analysis of health effects from population exposure to acid precipitation." Environmental Health Perspective **63**(45).

Parry, G. W. (1996). "The characterization of uncertainty in probabilistic risk assessments of complex systems." Reliability Engineering and System Safety **54**: 119-126.

Perrow, C. (1984). Normal Accidents: Living with High-risk Technologies. New York, Basic Books.

Peterson, C. R. and L. R. Beach (1967). "Man as an intuitive statistician." Psychological Bulletin **68**: 29-46.

Plato (1948 translation of work written before 347 B.C.). Euthyphro, Apology, Crito, Phaedo *The Death Scene*. New York, Macmillan.

Powers, D. (2005). Defense-in-Depth (DID). J. D. Brewer. Albuquerque, NM, Dana Powers, Senior Scientist, Sandia National Laboratories**:** the best, simple description of DID.

Rasmussen, N. C. (1975). Reactor safety study: An assessment of accident risks in U.S. commercial nuclear power plants. WASH-1400 (NUREG-75/014), US Nuclear Regulatory Commission.

Reason, J. (1990). Human Error. Cambridge, UK, Cambridge University Press.

Reason, J. (1997). Managing the risks of organizational accidents. Brookfield, VT, Ashgate Publishing Company.

Redelmeier, D. A. and E. Shafir (1995). "Medical decision making in situations that offer multiple alternatives." Journal of the Amercian Medical Association **273**(4): 302-305.

Ross, S. M. (2000). Introduction to Probability Models. San Diego, CA, Academic Press.

Rotter, J. B. (1966). "Generalized expectancies for internal versus external control of reinforcement." Psychological Monographs **80**: 1-27.

Rumsfeld, D. H. (2002). DoD News Briefing - Secretary Rumsfeld and Gen. Myers. United States Department of Defense.

Saunders, P. A. H. and B. O. Wade (1983). Radiation and its control. Nuclear Power Technology: Volume 3, Nuclear Radiation. W. Marshall. Oxford, Clarendon Press.

Savage, L. J. (1971). "Elicitation of personal probabilities and expectations." Journal of the American Statistical Association **66**: 783-801.

Schlaifer, R. (1959). Probability and statistics for business economics. New York, McGraw-Hill.

Scientists, Union of Concerned (1977). The risks of nuclear power reactors. Cambridge, MA.
SEAB (1993). Earning public trust and confidence: requisites for managing radioactive wastes. Washington, DC, U.S. Department of Energy: Secretary of Energy Advisory Board, Task Force on Radioactive Waste Management.
Sheridan, T. B. (2002). Humans and automation: system design and research issues. New York, John Wiley & Sons, Inc.
Sheridan, T. B. and W. R. Ferrell (1974). Man-machine systems: Information, Control, and Decision Models of Human Performance. Cambridge, MA, MIT Press.
Siegrist, M. and G. Cvetkovich (2001). "Better negative than positive? Evidence of a bias for negative information about possible health dangers." Risk Analysis **21**(1): 199-206.
Simon, H. A. (1957). Models of Man: Social and Rational. New York, Wiley.
Simon, H. A. and R. K. Sumner (1968). Patterns in Music. Formal Representation of Human Judgement. B. Kleinmuntz. New York, NY, John Wiley & Sons, Inc.**:** 219-250.
Slovic, P. (1966). "Cue consistency and cue utilization in judgment." American Journal of Psychology **79**: 427-434.
Slovic, P. (1986). "Informing and educating the public about risk." Risk Analysis **6**: 403-416.
Slovic, P. (1987). "Perception of risk." Science **236**: 280-285.
Slovic, P. (2000). The Perception of Risk. Sterling, VA, Earthscan.
Slovic, P., B. Fischhoff, et al. (1981). "Perceived Risk: Psychological Factors and Social Implications." Proceedings of the Royal Society of London **A 376**: 17-34.
Slovic, P., B. Fischhoff, et al. (1982). Response mode framing, and information-processing effects in risk assessment. New directions for methodology of social and behavioral science: Question framing andd response consistency. R. Hogarth. San Francisco, Jossey-Bass**:** 21-36.
Sober, E. (1981). "The principle of parsimoney." British Journal for the Philosophy of Science **32**: 145-156.
Sober, E. (1990). "Let's razor Ockham's Razor." Philosophy supplements: 73-93.
Starr, C. (1969). "Social benefit versus technological risk." Science **165**: 1232-1238.
Stevens, S.S. (1946). "On the theory of scales of measurement." Science **103**: 677-680.
Talty, J. T. (1988). Industrial Hygiene Engineering: Recognition, Measurement, Evaluation, and Control. New Jersey, Noyes Data Corporation.
Tenner, E. (1996). Why Things Bite Back: Technology and the Revenge of Unintended Consequences. New York, NY, Vintage Books.
Thaler, R. (1980). "Toward a positive theory of consumer choice." Journal of Economic Behavior and Organization **1**: 39-60.
Thompson, M. (2005). Are these towers safe? Why America's nuclear power plants are still so vulnerable to terrorist attack - and how to make them safer. A special investigation. Time**:** 34-38 & 43-48.
Tverskey, A. and D. Kahneman (1973). "Availability, a heuristic for judging frequency and probability." Cognitive Psychology **5**: 207-232.
Tversky, A. (1977). On the elicitation of preferences: Descriptive and prescriptive considerations. Conflicting objectives in decisions. D. Bell, R. L. Kenney and H. Raiffa. New York, Wiley**:** 209-222.

Tversky, A. and D. Kahneman (1974). "Judgment under uncertainty: Heuristics and Biases." Science **185**(4157): 1124-1131.
Tversky, A. and D. Kahneman (1981). "The framing of decisions and the psychology of choice." Science **211**(4481): 453-458.
Tversky, A. and D. Kahneman (1992). "Advances in Prospect Theory: Cumulative representation of uncertainty." Journal of Risk and Uncertainty **5**(4): 297-323.
Tversky, A. and D. J. Koehler (1994). "Support theory: A nonextensional representation of subjective probability." Psychological Review **101**(4): 547-567.
UIC (2004). Energy subsidies and external costs. Melbourne, Australia, Uranium Information Centre.
UNSC - United Nations Scientific Committee on Effects of Atomic Radiation (1988). Sources, effects, and risks of ionizing radiation. New York, NY, United Nations.
USNAS - U.S. National Academy of Sciences Committee on Biological Effects of Ionizing Radiation, (1990). Health effects of exposures to low levels of ionizing radiation. Washington, DC.
Voelz, G. L. (1975). Radiological health in occupational medicine practice. Occupational Medicine: Principles & Practical Applications. C. Zenz. Chicago, Year Book Medical Publishers.
Walker, J. S. (2004). A Nuclear Crisis in Historical Perspective: Three Mile Island. Berkeley, CA, University of California Press.
WHOI - Woods Hole Oceanographic Institute (2001). In Memoriam: Arnold B. Arons. http://www.whoi.edu/mr/obit.do?id=774, Woods Hole Oceanographic Institute.
Wickens, C. D. and J. Hollands (2000). Engineering psychology and human performance. Upper Saddle River, NJ, Prentice Hall.
Wikipedia (2005). Occam's Razor - Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Occam's_Razor, Wikipedia**:** 9.
Wilson, R., S. D. Colome, et al. (1980). Health Effects of Fossil Fuel Burning. Cambridge, MA, Ballinger.
Winterfeldt, D. V. and W. Edwards (1986). Decision Analysis and Behavioral Research. New York, Cambridge University Press.
Zio, E. and G. E. Apostolakis (1996). "Two methods for the structured assessment of model uncertainty by experts in performance assessments of radioactive waste repositories." Reliability Engineering and System Safety **54**: 225-241.

Distribution

22	MS0748	Jeffrey D. Brewer	6861
1	MS0188	D. Chavez, LDRD Office	1030
2	MS9018	Central Technical Files	8945-1
2	MS0899	Technical Library	4536

Second Printing, January 2006

92	MS0748	Jeffrey D. Brewer	6861

Sandia National Laboratories