

# MANAGING UNCERTAINTY IN ENGINEERING DESIGN USING IMPRECISE PROBABILITIES AND PRINCIPLES OF INFORMATION ECONOMICS

## PH.D. THESIS PROPOSAL

JASON AUGHENBAUGH  
AUGUST 31, 2005

## TABLE OF CONTENTS

<b>1</b>	<b>Introduction.....</b>	<b>1</b>
1.1	Goal of the Research.....	1
1.2	Background and Motivation.....	1
1.3	Summary of Research Plan.....	3
<b>2</b>	<b>Research Questions and Tasks .....</b>	<b>3</b>
2.1	Uncertainty in Engineering Design.....	4
2.1.1	Task 1: Identifying Characteristics of Uncertainty in Design.....	5
2.1.2	Justification of Hypothesis 1.....	5
2.2	Practical Value Assessments of Design Methods.....	8
2.2.1	Task 2: Creating a method to compare practical value.....	9
2.3	Managing Information Collection: Information Economics in Engineering Design ...	10
2.3.1	Task 3: Bounding the value of additional information collection.....	12
2.4	Decision-making with indeterminate information.....	12
2.4.1	Task 4: Identifying and characterizing decision policies and meta-data.....	12
<b>3</b>	<b>Validation and Example Problems.....</b>	<b>13</b>
3.1	Example Problem 1: Pressure Vessel Design.....	13
3.2	Example Problem 2: EBDM Example.....	13
<b>4</b>	<b>Schedule.....</b>	<b>14</b>
<b>5</b>	<b>Impact and Significance of Proposed Research.....</b>	<b>15</b>
	<b>References.....</b>	<b>16</b>

## SUMMARY

The engineering design community recognizes that an essential part of the design process is decision-making. Each decision consists of two main phases—problem formulation and problem solution. Existing literature focuses on problem *solution* using precisely known probabilities. Problem *formulation* has received considerably less attention. The objective of this thesis is to investigate methods for managing uncertainty during the formulation phase of engineering design decisions, focusing on situations in which probabilities are not precisely known. The existence of such situations has been recognized in the decision theory community but has not been addressed substantially in engineering design problems. The thesis seeks to identify the fundamental characteristics of uncertainty in the context of engineering design, and hypothesizes that subjective, *imprecise probabilities* are more general and appropriate than currently used representations. Many methods of uncertainty representation are rigorous and internally consistent, but the applicability of the starting axioms must be evaluated. Further, the thesis will develop a method for comparing the *practical value* of alternative problem formulations and uncertainty representations, taking into account not only the mathematical expressiveness of the formalism, but also the cost of the computations. This is an *information management* problem to which principles of *information economics* will be applied for determining an appropriate cost-benefit tradeoff. Finally, the thesis will evaluate decision policies for problems with imprecisely known probabilities.

# 1 Introduction

## 1.1 Goal of the Research

During the engineering design process, resources are limited, and hence the available information for making decisions is incomplete. Existing work has focused on solving design problems, both in terms of decision methods and computer models. The problem formulation has received much less attention. As computing power has increased, computations have become more precise and engineering models have become more complex, but the underlying models of uncertainty have remained the same—precise probability distributions. The goal of this thesis is to investigate methods for managing uncertainty during the formulation phase of engineering design decisions, focusing on situations in which probabilities are not precisely known. The main research question is: *How should engineering designers manage information to support decision making under uncertainty?* This question will be answered from two perspectives:

1. A *theoretical* perspective—identifying the internal consistency and applicability of representations to design problems
2. A *practical* perspective—given the benefits and costs associated with different methods, identifying the method that yields the highest overall economic value to the design process.

## 1.2 Background and Motivation

Engineering design is a hierarchical and iterative process, consisting of the phases of product planning, clarification of task, conceptual design, embodiment design, and detail design (Pahl and Beitz 1996). The decision-based design research community has formalized this process by focusing on decision making as the critical points in the design process (Hazelrigg 1996; Hazelrigg 1998; Marston, Allen et al. 2000). These decisions have two phases—problem *formulation* and problem *solution*. The decision formulation phase involves an important sub-decision problem, namely, how much information to collect in support of the decision making.

The basic elements of formulating a decision problem and collecting information can be summarized in five steps (Kmietowicz and Pearman 1981):

1. Identify an exhaustive set of mutually exclusive decision alternatives
2. Identify an exhaustive set of mutually exclusive alternative states of the world
3. Predict the payoff of every decision alternative in every state of the world
4. Assign probabilities to each state of the world, or acknowledge that such knowledge is not available
5. Select the criteria for evaluating alternatives

After formulating the problem, engineers can solve the problem by identifying the decision alternative that is most preferred. In general, engineers must complete the formulation process within significant constraints, such as deadlines, monetary budgets, and *bounded rationality*—a phrase coined by Herbert Simon (1947) that refers to the inherent bounds on human thinking. At a higher level, thinking takes time and energy, which, like all resources, are in short supply.

Despite such constraints on the process, the problem formulation phase of the decision problem has received much less attention than the problem solution phase. Recent literature (Bradley and Agogino 1994; Gupta and Xu 2002; Radhakrishnan and McAdams 2005) acknowledges the reality of resource constraints and the impracticality of exhaustive analysis, but it presents few

alternatives for problem formulation. Gupta and Xu (2002) note the impossibility, given time and budgetary constraints, of exploring all possible design alternative payoffs (steps 1-3), and they identify significant tradeoffs in the number of alternatives considered, but they present no guidance for actually managing the design process in this aspect. Radhakrishnan and McAdams (2005) analyze the cost-benefit trade-offs in selecting models of various levels of abstraction in engineering design and present a framework in which a designer can reason about model uncertainty, but the designer is left with little guidance in estimating actual value of information from different models. Bradley and Agogino (1994) develop a decision-analytic approach to assist designers in cost-benefit analysis of resource expenditures using precisely characterized probability distributions to guide and prioritize information collection, but they do not explain how to estimate these distributions.

Other work has focused on evaluation, Step 5. Engineers have developed or adopted various methods to support design decisions under uncertainty, such as statistical decision theory (Pratt, Raiffa et al. 1995), utility theory (von Neumann and Morgenstern 1980), the method of imprecision (Antonsson and Otto 1995), safety factors (Elishakoff 2004), probabilistic risk assessments (Bedford and Cooke 2001), reliability based design optimization (Mourelatos and Liang 2004), and robust design (Byrne and Taguchi 1987; Taguchi 1987). Each of these methods requires the designer to formalize preferences in some way. The process of eliciting and formalizing preferences is not necessarily trivial, but it is not the focus of this research. In this proposed dissertation, it is assumed that designers can capture their preferences in a utility function. The use of this utility function in decision making is addressed throughout this proposal.

A lingering assumption in most methods is the existence of clearly defined probabilities. That is, in most engineering design methods, it is assumed that Step 4 can be completed with little difficulty; the possibility that the probabilities are unknown is often ignored. From where do designers get these probabilities? How should these probabilities be represented? Are probabilities even appropriate for representing uncertainty in engineering design? These questions are exciting from both a philosophical perspective—what is uncertainty?—and from a practical perspective—what method is most valuable to the designer?

While philosophical questions are interesting to ponder, this thesis touches on them only briefly before tackling the practical issues—in practice, how should engineers overcome resource constraints and uncertainty in the design process? I plan to answer this question not only from a mathematical perspective, but also from an *economic* perspective. Different formalisms for uncertainty have different assumptions, different methods, and different costs associated with their use. Part of managing uncertainty in engineering design is choosing the “right” representation. Another part is collecting the “right” amount of information to make a decision. So how can “right” be defined? Extending the principles of *information economics* (Marschak 1974), an engineer should choose the methods, models, and information that accord the greatest net value—benefit minus cost—to the entire design process and product lifecycle. The implementation of these principles in uncertainty management could revolutionize the way engineers think about design. By focusing on the practical value, the usefulness of the methods can be evaluated directly. The stated questions and perspectives lead directly to the proposed primary research question for this dissertation:

<b>Primary Research Question:</b> How should engineering designers manage information to support decision making under uncertainty?
-------------------------------------------------------------------------------------------------------------------------------------

### 1.3 Summary of Research Plan

The proposed research plan divides the thesis into four sub-research questions and tasks, which are summarized in Figure 1 and explained in detail in the next section. An overview of the plan is as follows. The first part of the thesis will identify both the characteristics of uncertainty in engineering design and potential representations for this uncertainty. Given that different approaches exist, each rigorous and self-consistent but based on different axioms, a method needs to be developed by which the practical value of using one approach compared to the other is determined. Therefore, the second part of the thesis will present a method for comparing this practical value. In addition to choosing an appropriate representation of uncertainty, engineers must decide how much information to collect, given their limited resources. The third part of the thesis will explain the principles of information economics and present a method for estimating the value of collecting additional information in engineering design problems. Finally, engineers must be able to make decisions, so the fourth part of the thesis will determine and apply appropriate decision methods for the selected uncertainty representation.

## 2 Research Questions and Tasks

This section explains in detail the research questions, hypotheses, and tasks of the research plan outlined in the previous section and Figure 1. The section is organized by research tasks, since there is a one-to-one mapping between questions and tasks. The hypotheses will be validated in part through the use of two example problems: a pressure vessel design problem and an application in environmentally benign design and manufacture (EBDM). The details of the validation process and example problems are presented in Section 3.

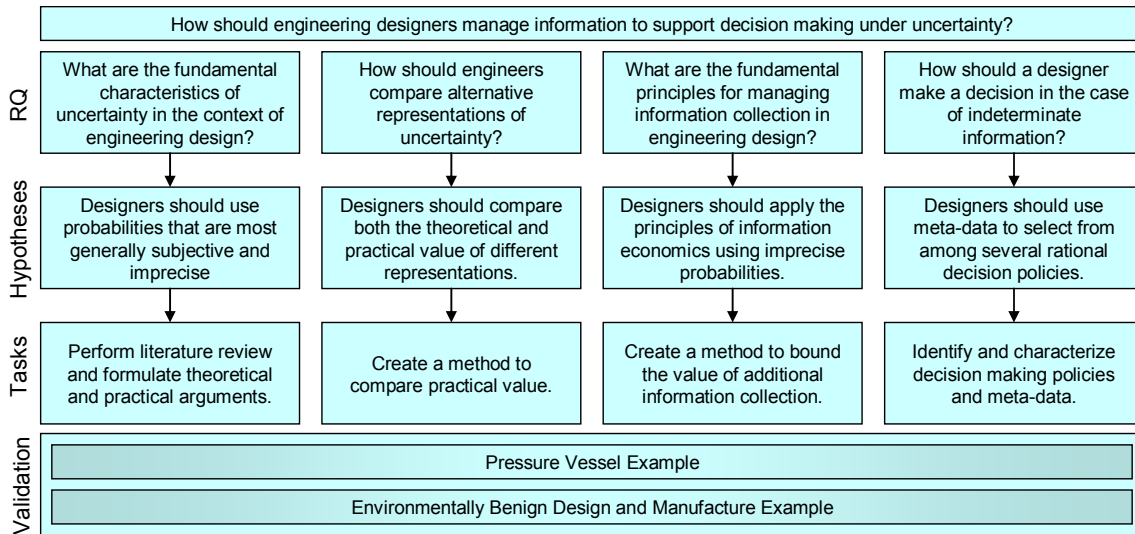


Figure 1: Research Overview

## 2.1 Uncertainty in Engineering Design

**Research Question 1:** What are the fundamental characteristics of uncertainty in the context of engineering design?

### Problem Characteristics:

In engineering, *uncertainty* is defined simply as a lack of certainty. *Certainty* is the condition of knowing everything—including future states of the world—necessary to choose the course of action whose outcome is most preferred (Nikolaidis 2005). A decision-maker's uncertainty is the gap between certainty and the information the decision-maker currently has available for decision making, which is the *present state* of information.

Decision-making under uncertainty has been studied for many years in a variety of contexts, including engineering design. In traditional statistical decision theory (Pratt, Raiffa et al. 1995), utility analysis, as originally proposed by von Neumann and Morgenstern (1980), is used for decision making. Utility theory has been studied extensively by economists and decision theorists, and there continues to be an increasing interest in applying utility methods to engineering problems (Thurston 1990; Hazelrigg 1998; Fernández, Seepersad et al. 2001; Scott 2004). In general, *utility* expresses preference—more preferred decision outcomes are assigned higher utility values. If chosen correctly, utilities reflect the decision-maker's preferences, even under uncertainty, as follows.

In utility theory, the likelihood of a state of the world occurring is expressed as a probability. The expected utility of a particular alternative is the mathematical expectation of the utility of that alternative, meaning the utility under each state of the world is weighted by that state's probability of occurring. According to utility theory, a rational designer would choose the optimal design alternative with the greatest expected utility.

Utility theory assumes that these probabilities are precisely known, or can effectively be estimated. However, decision theory dating back to at least Knight (1921) recognized that sometimes probabilities are not known. Other work has noted cases of incomplete knowledge of the probabilities, (Cannon and Kmietowicz 1974), such as an ordering or ranking of probabilities (Fishburn 1964). Other researchers have introduced the notion of probabilities that can only be bounded by upper and lower (or *imprecise*) probabilities (Good 1950; Smith 1961; Smith 1965; Good 1983; Kyburg 1987; Walley 1991; Voorbraak 1993).

In addition to different levels of knowledge of probabilities, there is an even more fundamental issue—what is a probability? There are many possible interpretations of probability (Savage 1972; de Finetti 1974; Walley 1991; Hajek 2003; Joslyn and Booker 2005), and at times the design community fails to adequately specify which interpretation they are using. Other recent work has even suggested the existence and necessary consideration of fundamentally different types of uncertainties (Smets 1991; Helton 1994; Ferson and Ginzburg 1996; Winkler 1996; Oberkampf, Helton et al. 2002; Berleant, Cheong et al. 2003; Dai, Scott et al. 2003; Nikolaidis 2005). Given the existence of differing interpretations of probability, which is appropriate for which application? Given limited resources, can probabilities ever be known precisely? Are there fundamentally different types of uncertainty that need to be considered in engineering design? Specifically, as an answer to Research Question 1, I advance the following hypothesis:

**Hypothesis 1:** Uncertainty in decision making is epistemic in nature, and thus engineering designers should use probabilities that are most generally subjective and imprecise, such as those that can be represented by probability-boxes.

### 2.1.1 Task 1: Identifying Characteristics of Uncertainty in Design

This task is divided into two subtasks, as follows.

#### TASK 1.1: LITERATURE SURVEY OF UNCERTAINTY

**Aim:** The objective is to identify the fundamental characteristics of uncertainty in engineering design and to select an appropriate mathematic formalism for such uncertainties.

**Method:** This task will be completed by performing an extensive literature survey of existing work in uncertainty, probability, and decision theory. Primary emphasis will be on demonstrating the internal consistency of the methods, their consistency with notions of rationality, and their connection to the practical realities of engineering design.

**Status:** The majority of this task has been completed over the past year, including preparation of a report for NASA and two conference papers (Aughenbaugh, Ling et al. 2005; Aughenbaugh and Paredis 2005).

#### TASK 1.2: THEORETICAL AND PRACTICAL ARGUMENTS

**Aim:** The objective is to create a coherent, logical argument for the use of subjective, imprecise probabilities and probability-boxes in engineering design based on strong axioms and specific requirements for an engineering design method.

**Method:** The task will be completed by creating strong theoretical arguments for subjective, imprecise probabilities including simple examples, as well as pursuing example design problems using probability-boxes to represent uncertainty. The comparison of the practical value of different methods will be addressed in Task 2.

**Status:** Probability-boxes have been used in the example design problem of a simple pressure vessel (Aughenbaugh and Paredis 2005), but has yet to be expanded to a EBDM example.

### 2.1.2 Justification of Hypothesis 1

Because Hypothesis 1—the use of subjective, imprecise probabilities—is relatively fundamental to the remainder of the proposal, this section expands on the argument for subjective, imprecise probabilities in an effort to demonstrate the value and feasibility of using them in engineering design.

#### The need for subjective probabilities

The variety of interpretations of probabilities and means for representing knowledge about probabilities provides many options for engineering design. In other research domains, the debates over interpretations of probabilities can be quite contentious. However, the engineering design literature appears relatively sparse on the matter. The most commonly adopted interpretations of probabilities in engineering design are variations of the *frequentist* and *subjective* interpretations (Joslyn and Booker 2005). I hypothesize that engineers should focus on a subjective interpretation because a frequentist interpretation is inherently impractical.

The *frequentist* interpretation is based on the notion of relative frequencies of outcomes. Under a frequentist interpretation, a probability represents the ratio of times that one outcome occurs compared to the total number of outcomes in a series of identical, repeatable, and possibly random trials (Reichenbach 1949; von Mises 1957). This presents several problems in engineering design. First, events are not always repeatable. In the extreme case of a one-time

event (such as a presidential election), the notion of relative frequencies is inadmissible—the outcome will either occur or not occur, exactly once. This makes standard utility theory inapplicable. Second, the assumption of *identical* trials is very strong, as it implies that the observed frequencies relate to exactly the same process, yet most experimental data comes from similar processes or prototypes, since decisions must be made before the final design is available for testing. Even assuming some events are essentially repeatable and data can be collected, a second problem is encountered: there is no guarantee that a particular sample is representative of the true relative frequency. Although in theory the relative sample frequency approaches the true relative frequency as the sample size goes to infinity, an infinite sample size is impossible in practice. Consequently, the true relative frequency can never be determined.

Proponents of a *subjective* interpretation of probability assert that there is no such thing as a true or objective probability, but rather probabilities are an expression of belief, as reflected in an individual's willingness to bet (de Finetti 1980; Winkler 1996; Joslyn and Booker 2005). When framed appropriately, such bets can be taken as subjective probabilities (de Finetti 1980). This interpretation allows engineers to express probabilities for one-time events. For example, the probability that market demand for a product will exceed one million dollars has no real meaning under a frequentist interpretation, because that event—the actual market demand—will occur exactly once. A subjective interpretation, on the other hand, allows a designer to express his or her belief as a probability, and then to apply expected utility maximization (Savage 1972) or another decision policy.

It seems that in practical problems, a subjective interpretation of probability is superior to a frequentist interpretation. However, there are other interpretations to consider, such as the classical (Keynes 1962) and logical (Carnap 1950; Popper 1959; Keynes 1962) interpretations, as well as questions of how to construct subjective probabilities that are reasonably consistent with the available information, as in the *rationalist* interpretation of subjective probabilities (Walley 1991). Additional research on these methods will be completed as part of the overall literature review.

### **Justifications for imprecise probabilities**

Preliminary research also indicates that engineers' knowledge of subjective probabilities is often incomplete. How can they not know their own beliefs precisely? An individual's subjective probabilities may remain imprecisely characterized for a variety of reasons (Walley 1991), including:

- Available evidence is incomplete or indeterminate
- Available information is inconsistent or conflicts
- Available information is of limited relevance, such as from a different but similar process

There are also practical reasons that an individual may want to express imprecise probabilities. Specifically, these relate to a cost-benefit analysis. Collecting evidence, such as statistical information, is expensive. Similarly, the constraints of bounded rationality mean that a "cost of thinking" is incurred when assessing and eliciting one's beliefs. Imprecision can thus result from both an unintentional or deliberate lack of accurate or sufficient introspection and reasoning.

A strong formalization and axiomization of *imprecise probabilities* is presented by Walley (1991), and the value of using imprecise probabilities in certain engineering design decisions has been demonstrated in preliminary work for this proposed dissertation (Aughenbaugh and Paredis 2005). A full discussion of Walley's formalization of imprecise probabilities (Walley 1991) is well outside the scope of this proposal, but I summarize a few essentials.

As noted previously, under a subjective interpretation, a probability represents an individual's willingness to enter a bet. Imprecise probabilities are essentially intervals of probabilities, with an associated upper and lower probability. More generally, probabilities are a special form of *previsions* (de Finetti 1974). A lower prevision represents a price at which you are sure you would buy a bet, and the upper prevision represents a rate at which you are sure you would buy the opposite of the bet. If the two are equal, then they jointly represent your *fair price* (de Finetti 1974) for the bet, the price at which you are willing to take either side of the bet. At a price between your upper and lower previsions, you are not willing to enter the bet on either side, because given your current state of information, you are not sure how betting at these prices will affect your expected payoff.

Many people subscribe to what Walley (1991) calls a *dogma of precision* that only precise probabilities are admissible, claiming that imprecise probabilities violate rationality. It is generally accepted that in order to be rational, one's probabilities must avoid a sure loss, a situation often referred to as a Dutch Book (de Finetti 1974; Walley 1991; Hajek 2003). The general idea is that if an individual's subjective probabilities violate certain rules, a group of bets can be constructed, all of which the individual is willing to accept, but the combination of which results in a sure loss; the person will lose money under any outcome. This argument is often presented in favor of precise probabilities and the axioms of Kolmogorov (1956) or de Finetti (1980). However, Walley (1991) presents axioms of coherence for imprecise probabilities that also avoid sure loss, thus negating this objection.

### The use of probability boxes

Developing the theoretical, practical, and philosophical arguments behind an interpretation of probability is a necessary but not sufficient step towards decision making under uncertainty. It is also necessary to establish a mathematical formalism for representing and manipulating such probabilities. Some alternatives to precise probabilities, such as fuzzy sets (Zadeh 1965), possibility theory (Zadeh 1978), fuzzy measures (Sugeno 1974; Sugeno 1977), the transferable belief model (Smets 1990), and the original interpretation of Dempster-Shafer Evidence Theory (Dempster 1967; Shafer 1976; Shafer 1992; Yager, Kacprzyk et al. 1994), abandon probability completely. This opens them up to significant criticism, at least in part due to their unfamiliarity to practicing engineers and lack of a clear numerical interpretation. Imprecise probabilities, as the name implies, are still probabilities, though obviously they are in some sense non-standard probabilities.

The recently developed formalism called probability bounds analysis (PBA) (Ferson and Donald 1998; Tucker and Ferson 2003) extends traditional probability theory and uses a structure called a *probability-box*, or *p-box*, to represent imprecise probability distributions, as shown in Figure 2. PBA has appealing properties, such as its ability to propagate uncertainty in a computationally efficient fashion. In addition, preliminary work for this proposed dissertation has demonstrated the value of using p-boxes to represent imprecise probabilities in an example pressure vessel design problem (Aughenbaugh and Paredis 2005).

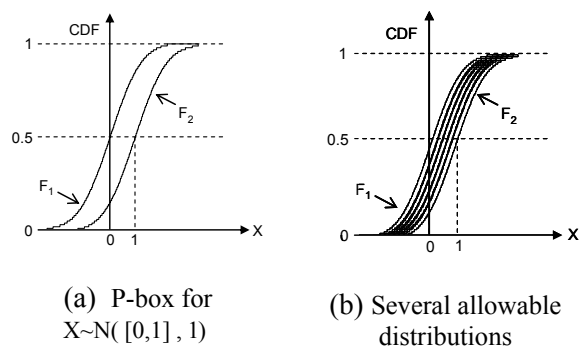


Figure 2: Example P-box and distributions

A p-box expresses interval bounds on the cumulative probability distribution function (CDF) for a random variable. In summary, a p-box is a more expressive generalization of both traditional probability theory and interval representations. A general p-box expresses both probability (represented in the shapes of the CDFs) and imprecision (represented by the separation between the upper and lower bounds). This makes PBA an ideal candidate for representing imprecise probabilities. More formally, the bounds on a p-box, such as shown in Figure 2(a), are given by two CDFs ( $F_1$  and  $F_2$ ) that enclose a set of CDFs (Figure 2(b)) that are, under some interpretation, consistent with the current state of information. Any distribution that falls partially or entirely outside of the p-box is definitely inconsistent with the present state of information.

## 2.2 Practical Value Assessments of Design Methods

**Research Question 2:** How should engineers compare alternative representations of uncertainty?

### Problem Characteristics:

Preliminary work towards completing Task 1 has revealed that there are multiple representations of uncertainty. If we think of a representation as a model of uncertainty, then from an engineering perspective, we can say that none of them is the “right” model of uncertainty—they are all abstractions.

No representation contains exactly the available information; the information state is modified in two directions. First, information is usually lost when abstracting to the representation. Second, information is often added. For example, consider a formalism for uncertainty that only allows point estimates to be stated. When representing information about tomorrow’s temperature in this formalism, any information regarding probability distributions or confidence levels is thrown away. At the same, the outcome now appears certain, since there is only one number representing the possible outcome. Most people know that this prediction is inexact and non-deterministic, but this meta-information is not captured in the formalism itself.

Each model begins from a particular set of axioms, and because the acceptance of a set of axioms is often a philosophical issue open to endless theoretical debate, logical reasoning cannot lead to a judgment of a “best” method. However, one can examine if there is a consistent interpretation of the mathematical axioms that relates them to reality.

Even if a particular formalism is clearly superior to another, such as more expressive mathematically than alternatives, it does not necessarily make it better from an engineering perspective. When choosing a representation for application in engineering design, we really need to adopt a more pragmatic perspective. An engineer needs to determine the value—benefit minus cost—of using different methods of uncertainty representation for a particular design problem. This is equivalent to selecting an appropriate model and level of abstraction for any engineering problem; sometimes it just is not valuable to model every nut and bolt of the system. We thus need a method for comparing the value of using different representations of uncertainty in engineering design practice.

**Hypothesis 2:** Representations of uncertainty must be shown to be internally consistent and then compared on the basis of the average value of using one representation over another in a particular *class* of engineering design problems, where the class is determined by considering meta-information not captured in the formalism.

## 2.2.1 Task 2: Creating a method to compare practical value

**Aim:** The objective is to develop and demonstrate a method for comparing the practical value of using different uncertainty representations in engineering design.

**Method:** This task will be completed by creating a method for comparing the practical value of using different uncertainty representations. The method will then be used to compare a subjective, imprecise probability model to alternative models for both the pressure vessel design example and the EBDM example.

**Status:** An approach has already been developed and demonstrated for the pressure vessel design example, but must be extended and applied to the EBDM example. This work has been described in a conference paper (Aughenbaugh and Paredis 2005). The method is summarized here.

The scenario is that a random manufacturing process is to be sampled, and a finite number of yield strength samples is presented to the designers. The experiment, shown in Figure 3, consists of two designers: one (approach A) using a single best-fit normal distribution to represent the uncertainty in the value of yield strength to design a pressure vessel, and the other (approach B) using a p-box to represent the uncertainty about the yield strength. Both approaches start with the same information about the uncertain yield strength—a set  $\Sigma$  of data samples. Both designers assume that the true distribution is Gaussian, but they use their own approach to represent their uncertainty about the distribution's parameters. Then they each choose an optimal design accordingly.

The goal of the experiment is to compare the utility of the design solutions that result from different approaches for representing uncertainty applied to the same design problem. The comparison is made possible in this experiment because we assume that overseeing the experiment is a supervisor who is in a state of precise information about the yield strength. From the supervisor's perspective, the uncertainty is precisely characterized by a normal distribution with a known mean  $\mu$  and a standard deviation of  $\sigma$ . In order to compare the value of the two approaches, the supervisor computes the difference in expected utility of the two designs, which is the value of approach B over approach A and can be expressed mathematically as

$$V(B) = E_{\sigma_y|\Sigma} [U(\sigma_y, b_\Sigma^*) - U(\sigma_y, a_\Sigma^*)],$$

where  $U(\sigma_y, b_\Sigma^*)$  is the utility of the design resulting from approach B given the realized

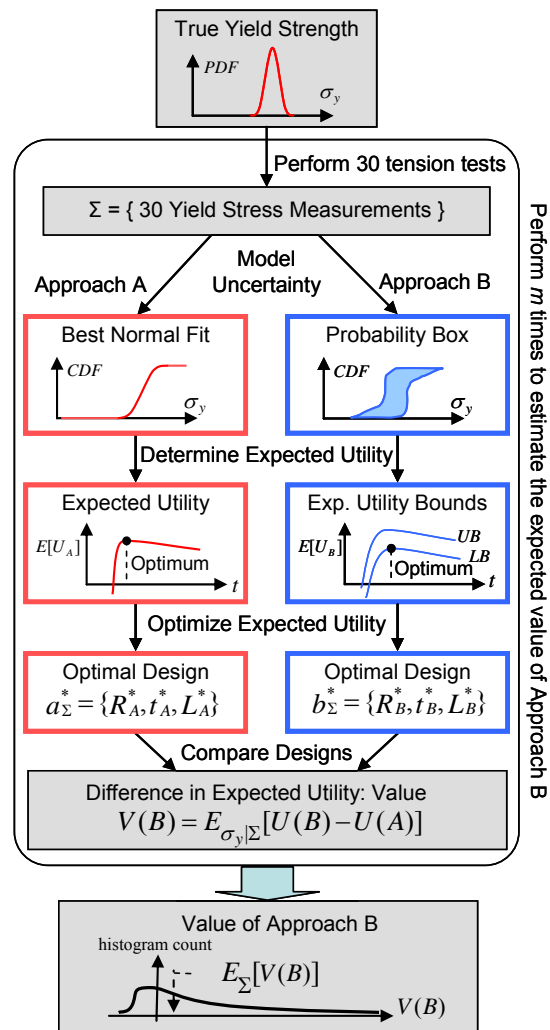


Figure 3: A computational experiment for determining the value of using imprecise probabilities

yield strength  $\sigma_y$  and  $U(\sigma_y, a_\Sigma^*)$  is the same but for approach A.

Because this value is for only one set of random data samples, one trial is not sufficient to judge the relative value of each approach; the supervisor needs to repeat the above experiment many times in order to determine which design approach performs best *on average*, over  $m$  different sample sets  $\Sigma$ . Mathematically, the expectation must be taken with respect to  $\Sigma$  in order to calculate the *average expected value of approach B over A*, written

$$E_\Sigma[V(B)] = E_\Sigma[E_{\sigma_y|\Sigma}[U(\sigma_y, b_\Sigma^*) - U(\sigma_y, a_\Sigma^*)]].$$

The addition of the word *average* emphasizes that this quantity is the expectation over the samples of the expected utility of particular design solutions. This value indicates to the supervisor whether approach B or approach A tends to perform better.

Preliminary results, shown in Figure 4, indicate that as the imprecision increases, the value of approach B over approach A increases significantly. As the imprecision approaches zero, approach B becomes only slightly worse than A. For different design problems, the designer will not necessarily know where the two curves cross. Thus, unless the designer is sure *a priori* that the consequences of the imprecision are insignificant, the results of this computational experiment suggest that it is valuable to explicitly represent the imprecision in the available characterization of uncertainties by using imprecise probabilities.

### 2.3 Managing Information Collection: Information Economics in Engineering Design

**Research Question 3:** What are the fundamental principles for managing information collection in engineering design?

#### Problem Characteristics:

Once probabilities are allowed to be subjective and imprecise, an obvious question arises: *how imprecise can probabilities be and still allow for good decision making?* This is a sub-question to a more general question: *how much information should a designer have before making a decision?*

The collection of information requires the expenditure of resources, whether this involves performing more experiments, building more prototypes, or creating additional simulations. At the same time, additional information can lead to better decisions, thereby increasing the payoff of a design. The existence of such cost-benefit tradeoffs in engineering design is obvious, but little research has been done into how to make these tradeoffs during the formulation phase of a decision problem. Research has instead focused on tradeoffs in the performance of the *final design*, rather than the design process. The process of collecting information to support decisions is resource intensive, and as such the costs of formulating design decisions impact the net payoff of the final design. As noted in the introduction, some work has recently explored tradeoffs in the

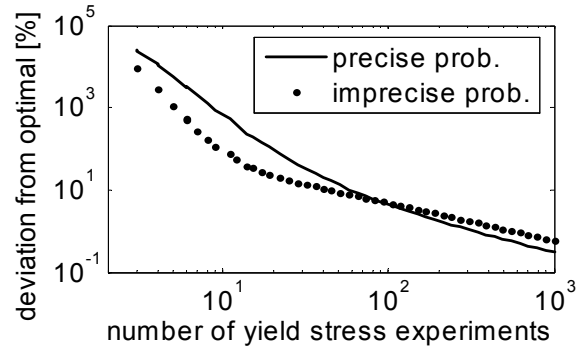


Figure 4: Variation of value with imprecision

formulation phase from a theoretical perspective, but this work does not present a formalized method for practically measuring these tradeoffs during the design process.

To answer this question of managing information collection and problem formulation, I turn to *economics*—the *study of choice under conditions of scarcity* (Lieberman and Hall 2000). Extending this definition, *information economics* is the study of choice in information collection and management when resources to expend on information collection are scarce. Designers face a scarcity of resources (identified above) in the design process; hence the principles of information economics can be applied to the information collection process in engineering design.

The area of information economics grew out of statistical decision theory in the 1950s when Marschak published a series of papers, consolidated in (Marschak 1974), on the economics of information and organization. Recently, with the explosion of new information technologies, information economics has regained attention within the broader context of information management (Rubin 1983; Strassmann 1999). Current areas of research focus on corporate finance and industry policy, such as intellectual property rights, industry regulation, and fostering innovation, or on the infusion of information technology into a corporation (Strassmann 1990; Strassmann 2004). Within engineering, the focus of information management has been primarily on data exchange, interoperability, and visualization to support collaborative design (for an overview of these areas, refer to the following review articles (Ciocoiu, Gruninger et al. 2001; Jayaram, Vance et al. 2001; Rangan and Chadha 2001; Szykman, Sriram et al. 2001; Urban, Dietrich et al. 2001)). In a more general sense, information economics presents principles by which the cost-benefit tradeoffs of information collection can be managed in engineering design. Many of these concepts have been developed and employed previously in standard micro-economics and the theory of the economic value of information, pioneered by Marschak (1974) and summarized by Lawrence (1999).

One way information can provide benefit is by leading to a decision action with a higher payoff. Let  $\pi(a, x)$  represent the payoff of action  $a$  given a state of the world  $x$ , and let  $a_y^*$  denote the designer's chosen decision action with additional information  $y$ , and  $a_0^*$  denote the designer's prior decision action before receiving information  $y$ . Finally let  $\text{cost}(y)$  denote the cost of acquiring  $y$ . Then the net value of a message is

$$v(x, y) = \pi(a_y^*, x) - \pi(a_0^*, x) - \text{cost}(y).$$

The complication enters when considering what will happen on average, since neither the state of the world  $x$  or the actual information  $y$  is known *a priori*. The existing work by Marschak (1974) and Lawrence (1999) assume that the joint probability density function  $p(x, y)$  is perfectly known, but this is rarely the case in engineering design. For example, the purpose of much information collection is to estimate  $p(x)$ , as in the yield stress test described in Task 2. It is also often the case that  $p(y)$  is unknown, such as in the case of a prototype construction. Does it really make sense to state precisely that the prototype design will yield a particular result (or message)  $y_i$  with probability  $p(y_i)$ ?

Given these limitations, do engineers have any hope of even bounding the value of information? I advance the following hypothesis.

**Hypothesis 3:** The value of additional information collection can be bounded by extending the principles of information economics to include imprecise probabilities and then applying these principles to the problem formulation.

### 2.3.1 Task 3: Bounding the value of additional information collection

*Aim:* The objective is to establish principles and develop a method for managing information collection during the problem formulation phase of design decisions.

*Method:* The task will be completed by extending the principles of information economics, as pioneered by Marschak (1974) and summarized by Lawrence (1999). The established principles and methods are only applicable when probabilities are precisely known. This framework will be extended to make use of imprecise probabilities, thus making the principles applicable to engineering design information collection. Finally, a method for bounding the value of information collection will be developed and demonstrated using the pressure vessel and EBDM examples.

*Status:* Preliminary research for this proposed dissertation has provided a method for bounding the value of information from statistical data sampling, as demonstrated for the pressure vessel design example in a conference paper (Aughenbaugh, Ling et al. 2005). The method and requisite vocabulary are too complicated to explain in the scope of this proposal, so the readers are referred to the paper.

## 2.4 Decision-making with indeterminate information

**Research Question 4:** How should a designer make a decision in the case of indeterminate information?

To this point, this proposal has focused on the formulation of the problem. In this section, the focus moves to problem solution with imprecise probabilities. Methods (such as utility theory) that assume precise probabilities provide a clear, unambiguous means for choosing an optimal solution; each alternative has an associated expected utility, and the designer chooses the alternative with the highest expected utility. But what happens when imprecise probabilities are used? Since the probabilities are imprecise, the calculated expected utilities are also imprecise and take the form of intervals. A comparison of two alternatives (say  $X$  and  $Y$ ) thus involves the comparison of two intervals. If there is no overlap in the intervals, then one solution is clearly better, since its lower-bound on expected utility exceeds the upper-bound of the other solution. But what happens when the intervals overlap?

If intervals overlap, several scenarios can occur. Given the two intervals, all that is known is that each true, but unknown value, lies somewhere in each interval. Depending on where the true expected utilities lie, design  $X$  could be preferred to  $Y$ ,  $Y$  preferred to  $X$ , or the decision maker could be indifferent. Given the intervals, the most preferred alternative is indeterminate. What can a designer do? I advance the following hypothesis:

**Hypothesis 4:** Given overlapping intervals of expected utility, any non-dominated decision can be defended as rational, though different options may perform better in certain scenarios. In these cases, meta-information about the scenario will lead to selection of the best decision policy.

### 2.4.1 Task 4: Identifying and characterizing decision policies and meta-data

*Aim:* The objective is to identify potential decision making policies, and then to explore the value of using different methods for different design problems in order to characterize which policy are the best practical choice for a class of design problems.

**Method:** The first part of task 4 is to identify possible methods for resolving indeterminacy. The next step is to identify the requirements for applying such policies in engineering design. This includes constructing a strong theoretical argument for the rationality of all of these methods. Essentially, it is necessary to establish that the existence of indeterminacy is acceptable, and that arbitrary resolution of indeterminacy is an admissible approach. The next part of task 4 is to adapt these methods for engineering design problems using imprecise probabilities, and to demonstrate their application in the two example problems. The results of the demonstrations will then be combined with the requirements to provide recommendations for appropriate methods under various scenarios of available information, designer preferences, and design problem classes.

**Status:** A preliminary literature review has been conducted, and contact has been made with several decision theory researchers in the imprecise probability community. This task is the biggest reach and largest risk in the proposal. At a minimum, the application of various policies will be formalized and demonstrated for engineering design, even if no strong recommendations can be made.

### **3 Validation and Example Problems**

In general, validation will be performed at both a theoretical level (demonstrating the internal consistency of the methods) and at a practical level (demonstration in two example problems). Theoretical validation will be achieved by presenting clearly stated axioms, developing consistent logical arguments, and connecting axioms to principles of rationality and economic theory. The proposed research also involves considering practical aspects of using imprecise probabilities and p-boxes in engineering design problems. This is explored via two example problems.

#### **3.1 Example Problem 1: Pressure Vessel Design**

In the first example problem, a pressure vessel is to be designed to meet certain requirements while maximizing utility, as described in an existing publication (Aughenbaugh and Paredis 2005). The complication is that the pressure vessel is to be built using a new type of steel, whose variations in yield strength have only been measured in a set  $\Sigma$  of  $n$  independent tension tests, where  $n$  is a relatively small number due cost considerations. The designers believe that the yield strength is well modeled as a normally distributed random variable, but with finite data, they cannot know the parameters of the normal distribution precisely. The exact assumptions of the example (such as the utility function and decision policies) can be varied to explore the performance over a range of scenarios. While the example is relatively simple, preliminary results for Tasks 2 (Aughenbaugh and Paredis 2005) and Task 3 (Aughenbaugh, Ling et al. 2005) indicate that the results can be very informative.

#### **3.2 Example Problem 2: EBDM Example**

The second design example will be taken from the application area of environmentally benign design and manufacture (EBDM). A key hurdle in many EBDM decision problems (which can span an entire product life-cycle from cradle to ultimate disposal) is a lack of information, and hence uncertainty. In general, the level of uncertainty varies with different phases of the lifecycle. While some uncertainty may be accurately characterized using precise probability distributions, other aspects of a product's performance, such as its end-of-life and disposal, are very distant in time, and uncertainty about them can be severe. It therefore appears that EBDM is an ideal area in which to explore the management of uncertainty and the application of imprecise probabilities. I propose to work closely with Dr. Chris Paredis and Dr. Bert Bras, who share an

NSF grant to perform research in this area, in developing this example. Dr. Bras is an expert in EBDM applications and has existing example problems that I can draw from and adapt.

## 4 Schedule

**Deliverables:** The proposed schedule is illustrated in Figure 5. The top part shows the major tasks, and the bottom part shows more specific tasks and deliverables. The schedule includes five papers, on four of which I am the primary author. The journal papers are based heavily on existing papers, but will require some extension and reworking. I am the primary author on one, and the second author on the other. One DETC-2006 paper will include preliminary work from task 4 (decision policies), and the other will be a joint paper with EBDM collaborators based on tasks 2 and 3. The final deliverables are the dissertation and presentation.

**Status:** Task 1 is nearly complete, and I have already published papers related to the pressure vessel example for Task 2 (Aughenbaugh and Paredis 2005) and Task 3 (Aughenbaugh, Ling et al. 2005), indicated as “Task 2: PV” and “Task 3: PV”. This leaves the EBDM example validation for tasks 2 and 3, as well as all of task 4.

**Risk areas:** It is possible that the promising preliminary results are applicable only in the narrow class of problems related to the pressure vessel example, and consequently will not scale to the EBDM example. The main backup plan is to shift the emphasis to what distinguishes EBDM from other engineering applications, what conditions must be met for imprecise probabilities to be valuable, and when engineers can and cannot make value of information estimates. Because existing work mostly has ignored the issue, findings that conclusively demonstrated the lack of value would be useful in their own right. There is also risk involved in task 4, of which the primary contribution is to demonstrate how different decision methods apply and perform in engineering design problems. It is possible that the best method depends on designer preferences, available information, and the specific design problem. In this case, the back up plan is to thoroughly analyze one scenario and reach a strong conclusion regarding relevant meta-information in that scenario, at the expense of generality.

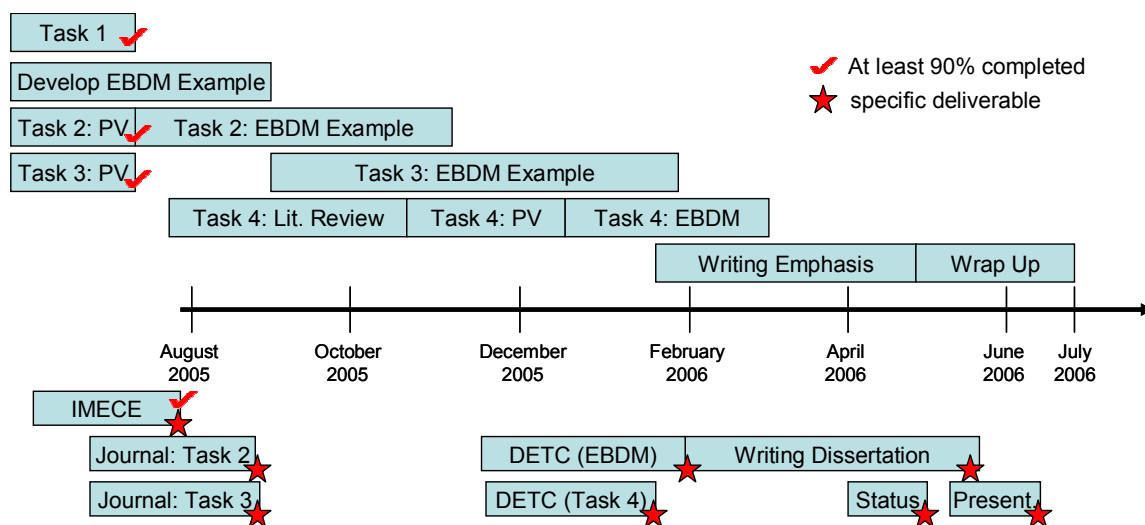


Figure 5: Proposed Schedule

## 5 Impact and Significance of Proposed Research

As a result of this research effort, primary principles for managing information and uncertainty in engineering design decisions will be formulated and demonstrated. A significant contribution is taking these principles out of the spheres of philosophical debate and placing them squarely in engineering design examples. The introduction of imprecise probabilities into the design process could significantly alter the paradigm of design, especially as it enables the active management of information during the problem formulation phase of design. Optimistically, this could significantly streamline the design process, saving development cost and improving product design. The specific contributions of this proposed research can be summarized as

1. Identifying the nature of uncertainty in engineering design
2. Determining under which circumstances imprecise probabilities are valuable
3. Creating a method for bounding the value of additional data collection
4. Identifying the relationship between meta-information about a decision problem and the practical choice of a rational decision policy
5. Establishing the practical value of using imprecise probabilities in two example problems

## References

- Antonsson, E. K. and K. N. Otto (1995). "Imprecision in Engineering Design." Journal of Mechanical Design **117**: 25-32.
- Aughenbaugh, J. M., J. M. Ling, et al. (2005). Applying Information Economics and Imprecise Probabilities to Data Collection in Design. 2005 ASME International Mechanical Engineering Congress and Exposition, paper no. IMECE2005-81181, Orlando, FL, USA.
- Aughenbaugh, J. M. and C. J. J. Paredis (2005). The Value of Using Imprecise Probabilities in Engineering Design. ASME DETC Design Theory and Methodology, Long Beach, CA, USA.
- Bedford, T. and R. M. Cooke (2001). Probabilistic risk analysis: foundations and methods. New York, Cambridge University Press.
- Berleant, D., M.-p. Cheong, et al. (2003). "Dependable handling of uncertainty." Reliable Computing **9**(6): 407-18.
- Bradley, S. R. and A. M. Agogino (1994). "An Intelligent Real Time Design Methodology for Component Selection: An Approach to Managing Uncertainty." Journal of Mechanical Design **116**(December): 980-988.
- Byrne, D. M. and S. Taguchi (1987). The Taguchi Approach to Parameter Design. Quality Congress.
- Cannon, C. M. and Z. W. Kmietowicz (1974). "Decision theory and incomplete knowledge." Journal of Management Science **11**(3).
- Carnap, R. (1950). Logical Foundations of Probability, Chicago University Press.
- Ciocoiu, M., M. Gruninger, et al. (2001). "Ontologies for Integrating Engineering Applications." ASME Journal of Computing and Information Science in Engineering **1**(1).
- Dai, Z., M. J. Scott, et al. (2003). Incorporating epistemic uncertainty in robust design. 2003 ASME Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Sep 2-6 2003, Chicago, IL, United States, American Society of Mechanical Engineers.
- de Finetti, B. (1974). Theory of Probability: A Critical Introductory Treatment. New York, Wiley.
- de Finetti, B. (1980). Foresight. Its Logical Laws, Its Subjective Sources (translated). Studies in Subjective Probability. H. E. Kyburg and H. E. Smokler, E. Krieger Publishing Company.
- Dempster, A. P. (1967). "Upper and Lower Probabilities Induced by a Multivalued Mapping." The Annals of Statistics **28**: 325-339.
- Elishakoff, I. (2004). Safety Factors and Reliability: Friends or Foes? Dordrecht, Netherlands, Kluwer Academic Publishers.
- Fernández, M. C., C. C. Seepersad, et al. (2001). Utility-Based Decision Support for Selection in Engineering Design. ASME DETC Design Automation Conference, Pittsburgh, PA.
- Ferson, S. and S. Donald (1998). Probability bounds analysis. International Conference on Probabilistic Safety Assessment and Management (PSAM4), New York, NY, Springer-Verlag.
- Ferson, S. and L. Ginzburg (1996). "Different Methods are Needed to Propagate Ignorance and Variability." Reliability Engineering and System Safety **54**(2-3).
- Fishburn, P. (1964). Decision and value theory. New York, J. Wiley.
- Good, I. J. (1950). Probability and the weighting of evidence, Hafner.
- Good, I. J. (1983). Good thinking: the foundations of probability and its applications. Minneapolis, University of Minnesota Press.
- Gupta, S. K. and C. Xu (2002). Estimating the Optimal Number of Alternatives to Be Explored in Large Design Spaces: A Step Towards Incorporation Decision Making Cost in Design Decision Models. ASME 2002 Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Montreal, Canada.
- Hajek, A. (2003). "Interpretations of Probability." The Stanford Encyclopedia of Philosophy Summer 2003. 2004, from <http://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=probability-interpret>.
- Hazelrigg, G. A. (1996). Systems Engineering: An Approach to Information-Based Design. Englewood Cliffs, NJ, Prentice Hall.
- Hazelrigg, G. A. (1998). "A Framework for Decision-Based Design." Journal of Mechanical Design **120**(4): 653-658.
- Helton, J. C. (1994). "Treatment of uncertainty in performance assessments for complex systems." Risk Analysis **14**: 483-511.

- Jayaram, S., J. Vance, et al. (2001). "Assessment of VR technology and its applications to engineering problems." ASME Journal of Computing and Information Science in Engineering **1**(1).
- Joslyn, C. A. and J. M. Booker (2005). Generalized Information Theory for Engineering Modeling and Simulation. Engineering Design Reliability Handbook. E. Nikolaidis, D. M. Ghiocel and S. Singhal. New York, CRC Press: 9.1-9.40.
- Keynes, J. M. (1962). A Treatise on Probability. New York, Harper and Row.
- Kmietowicz, Z. W. and A. D. Pearman (1981). Decision theory and incomplete knowledge. Aldershot, England, Gower.
- Knight, F. H. (1921). Risk, uncertainty and profit. Boston, Houghton Mifflin.
- Kolmogorov, A. N. (1956). Foundations of Probability. New York, Chelsea Publishing Company.
- Kyburg, H. E. (1987). Objective Probabilities. IJCAI-87.
- Lawrence, D. B. (1999). The Economic Value of Information. New York, Springer-Verlag.
- Lieberman, M. and R. Hall (2000). Introduction to Economics, South-Western College Publishing.
- Marschak, J. (1974). Economic Information, Decision, and Prediction: Selected Essays. Boston, D. Reidel Publishing Company.
- Marston, M., J. K. Allen, et al. (2000). "The Decision Support Problem Technique: Integrating Descriptive and Normative Approaches in Decision Based Design." Engineering Valuation and Cost Analysis **3**: 107-129.
- Mourelatos, Z. P. and J. Liang (2004). An Efficient Unified Approach for Reliability and Robustness in Engineering Design. NSF Workshop on Reliable Engineering Computing, Savannah, Georgia.
- Nikolaidis, E. (2005). Types of Uncertainty in Design Decision Making. Engineering Design Reliability Handbook. E. Nikolaidis, D. M. Ghiocel and S. Singhal. New York, CRC Press.
- Oberkampf, W., J. Helton, et al. (2002). "Epistemic Uncertainty Workshop." Reliable Computing **8**(6): 503-505.
- Pahl, G. and W. Beitz (1996). Engineering Design: A Systematic Approach. London, Springer Publishing.
- Popper, K. R. (1959). The Logic of Scientific Discovery. New York, Basic Books.
- Pratt, J. W., H. Raiffa, et al. (1995). Introduction to Statistical Decision Theory. Cambridge, MA, The MIT Press.
- Radhakrishnan, R. and D. A. McAdams (2005). "A Methodology for Model Selection in Engineering Design." Journal of Mechanical Design **127**(May): 378-387.
- Rangan, R. M. and B. Chadha (2001). "Engineering Information Management to Support Enterprise Business Processes." ASME Journal of Computing and Information Science in Engineering **1**(1): 32-40.
- Reichenbach, H. (1949). The Theory of Probability. Berkeley, University of California Press.
- Rubin, M. R. (1983). Information Economics and Policy in the U.S. Littleton, CO, USA, Libraries Unltd.
- Savage, L. J. (1972). The Foundation of Statistics. New York, Dover Publications.
- Scott, M. J. (2004). Utility Methods in Engineering Design. Engineering Design Reliability Handbook. New York, CRC Press.
- Shafer, G. (1976). A Mathematical Theory of Evidence. Princeton, NJ, Princeton University Press.
- Shafer, G. (1992). "Rejoinders to comments on "perspectives on the theory and practice of belief functions"." International Journal of Approximate Reasoning **6**(3): 445-480.
- Simon, H. (1947). Administrative Behavior. New York, The Macmillan Company.
- Smets, P. (1990). The transferable belief model: a model for quantified beliefs. IEE Colloquium on 'Reasoning Under Uncertainty' (Digest No.086), 22 May 1990, London, UK, IEE.
- Smets, P. (1991). "Varieties of ignorance and the need for well-founded theories." Information Sciences **57-58**: 135-144.
- Smith, C. A. B. (1961). "Consistency in statistical inference and decision." Journal of Royal Statistical Society **B23**: 1-37.
- Smith, C. A. B. (1965). "Personal Probabilities and statistical analysis." Journal of Royal Statistical Society **A128**: 469-499.
- Strassmann, P. A. (1990). The Business Value of Computers, The Information Economics Press.
- Strassmann, P. A. (1999). Information Productivity, Information Economics Press.
- Strassmann, P. A. (2004). "The Foundations of Information Economics - Part I." Information Economics Journal **March 2004**.
- Sugeno, M. (1974). The Theory of Fuzzy Integrals and Its Applications. Tokyo, Tokyo Institute of Technology.

- Sugeno, M. (1977). Fuzzy measures and fuzzy integrals: a survey. Fuzzy Automata and Decision Processes. M. M. Gupta, G. N. Saridis and B. R. Gaines.
- Szykman, S., R. Sriram, et al. (2001). "The Role of Knowledge in Next-generation Product Development Systems." ASME Journal of Computing and Information Science in Engineering **1**(1): 3-11.
- Taguchi, G. (1987). System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost. Dearborn, Michigan, UNIPUB/Kraus International Publications.
- Thurston, D. L. (1990). "Multiattribute utility analysis in design management." IEEE Transactions on Engineering Management **37**(4): 296-301.
- Tucker, W. T. and S. Ferson (2003). Probability bounds analysis in environmental risk assessment. Setauket, New York, Applied Biomathematics.
- Urban, S. D., S. W. Dietrich, et al. (2001). "Interconnection of Distributed Components: Current Middleware Solutions." ASME Journal of Computing and Information Science in Engineering **1**(1).
- von Mises, R. (1957). Probability, Statistics and Truth. London, Allen and Unwin.
- von Neumann, J. and O. Morgenstern (1980). Theory of games and economic behavior, third edition. Princeton, NJ, Princeton University Press.
- Voorbraak, F. (1993). As Far as I Know: Epistemic Logic and Uncertainty, Utrecht University.
- Walley, P. (1991). Statistical Reasoning with Imprecise Probabilities. New York, Chapman and Hall.
- Winkler, R. L. (1996). "Uncertainty in probabilistic risk assessment." Reliability Engineering & System Safety **54**(2-3): 127-132.
- Yager, R. R., J. Kacprzyk, et al. (1994). Advances in the Dempster-Shafer Theory of Evidence. New York, John Wiley and Sons.
- Zadeh, L. A. (1965). "Fuzzy Sets." Inf. Control **8**: 338-353.
- Zadeh, L. A. (1978). "Fuzzy sets as a basis for a theory of possibility." Fuzzy Sets and Systems **1**: 3-28.