1990-03

CERTS: a comparative evaluation method for risk management methodologies and tools

Garrabrants, William M.

Monterey, California. Naval Postgraduate School

(2)

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

CERTS: A Comparative Evaluation
Method for Risk Management Methodologies
and Tools

by

William M. Garrabrants
and
Alfred W. Ellis III

March 1990

Thesis Advisor: Lance J. Hoffman

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b RESTRICTIVE MARKINGS	
2a SECURITY CLASSIFICATION AUTHORITY		3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.	
2b DECLASSIFICATION/DOWNGRADING SCHEDULE			
4 PERFORMING ORGANIZATION REPORT NUMBER(S)		5 MONITORING ORGANIZATION REPORT NUMBER(S)	
6a NAME OF PERFORMING ORGANIZATION Naval Postgraduate School	6b OFFICE SYMBOL *(If applicable)* 55	7a NAME OF MONITORING ORGANIZATION Naval Postgraduate School	
6c ADDRESS *(City, State, and ZIP Code)* Monterey, CA 93943-5000		7b ADDRESS *(City, State, and ZIP Code)* Monterey, CA 93943-5000	
8a NAME OF FUNDING/SPONSORING ORGANIZATION	8b OFFICE SYMBOL *(If applicable)*	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	
8c ADDRESS *(City, State, and ZIP Code)*		10 SOURCE OF FUNDING NUMBERS	
		Program Element No	Project No / Task No / Work Unit Accession Number

11 TITLE *(Include Security Classification)*
CERTS: A Comparative Evaluation Method for Risk Management Methodologies and Tools

12 PERSONAL AUTHOR(S) Garrabrants, William M. and Ellis III, Alfred W.

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT *(year, month, day)*	15 PAGE COUNT
Master's Thesis	From To	March 1990	121

16 SUPPLEMENTARY NOTATION
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17 COSATI CODES			18 SUBJECT TERMS *(continue on reverse if necessary and identify by block number)*
FIELD	GROUP	SUBGROUP	Risk, Risk Management, Risk Assessment, Risk Analysis, Computer Security, Metrics

19 ABSTRACT *(continue on reverse if necessary and identify by block number)*

This thesis develops a comparative evaluation method for computer security risk management methodologies and tools. The subjective biases inherent to current comparison practices are reduced by measuring unique characteristics of computer security risk management methodologies. Standardized criteria are established and described by attributes which in turn are defined by metrics that measure the characteristics. The suitability of a method or tool to a particular organizational situation can then be analyzed objectively. Additionally, our evaluation method facilitates the comparison of methodologies and tools to each other. As a demonstration of its effectiveness, our method is applied to four distinct risk management methodologies and four risk management tools. Alternative models for utilizing the evaluation method are presented as well as possible directions for their application.

Without an adequate means of comparing and evaluating risk management decision-making methodologies, the metadecision (the selection of a risk management method or tool) becomes arbitrary and capricious, thereby making an inappropriate selection more likely. Selection of an inappropriate method or tool could lead to excessive costs, misdirected efforts, and the loss of assets. The systematic and standard comparison method developed in this thesis resolves that problem.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION	
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS REPORT ☐ DTIC USERS	Unclassified	
22a NAME OF RESPONSIBLE INDIVIDUAL Lance J. Hoffman	22b TELEPHONE *(Include Area code)* (202) 994-4955	22c OFFICE SYMBOL EE/CS

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

Approved for public release; distribution is unlimited.

CERTS: A Comparative Evaluation Method for Risk Management Methodologies and Tools

by

William M. Garrabrants
Major, United States Marine Corps
B.S., University of Washington, 1975

and

Alfred W. Ellis III
Major, United States Marine Corps
B.S. University of West Florida, 1984

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN INFORMATION SYSTEMS

from the

NAVAL POSTGRADUATE SCHOOL
March 1990

Authors:

[redacted]

William M. Garrabrants

[redacted]

Alfred W. Ellis III

Approved by:

[redacted]

Lance J. Hoffman, Thesis Advisor

[redacted]

Magdi Kamel, Second Reader

[redacted]

David R. Whipple, Chairman
Department of Administrative Sciences

ABSTRACT

This thesis develops a comparative evaluation method for computer security risk management methodologies and tools. The subjective biases inherent to current comparison practices are reduced by measuring unique characteristics of computer security risk management methodologies. Standardized criteria are established and described by attributes which in turn are defined by metrics that measure the characteristics. The suitability of a method or tool to a particular organizational situation can then be analyzed objectively. Additionally, our evaluation method facilitates the comparison of methodologies and tools to each other. As a demonstration of its effectiveness, our method is applied to four distinct risk management methodologies and four risk management tools. Alternative models for utilizing the evaluation method are presented as well as possible directions for their application.

Without an adequate means of comparing and evaluating risk management decision-making methodologies, the metadecision (the selection of a risk management method or tool) becomes arbitrary and capricious, thereby making an inappropriate selection more likely. Selection of an inappropriate method or tool could lead to excessive costs, misdirected efforts, and the loss of assets. The systematic and standard comparison method developed in this thesis resolves that problem.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial assistance and guidance provided by Dave Hsiao from the Naval Postgraduate School and Sylvan Pinsky from the National Computer Security Center. Stuart Katzke and Irene Gilbert from the National Institute of Standards and Technology were very supportive of our efforts and provided us the opportunity to experience the use of automated tools for computer security risk management.

Our thesis committee was magnificent. As our second reader, Magdi Kamel was enthusiastic with the project from inception to signature and had many excellent recommendations. Our advisor, Lance Hoffman, managed a long-distance relationship with us over the Defense Data Network with extraordinary finesse. His encouragement and enthusiasm were a constant source of inspiration. To each of these individuals, we extend a deep and heart-felt thank you.

Most importantly, we would like to express our deepest love and appreciation to our wives, Marie and Pam, for their interminable patience while this thesis became a reality. Without their understanding and support, we would probably still be writing.

I. INTRODUCTION

A. BACKGROUND

As the use of computers progressed in the early stages of the technology's development, the requirement for security was recognized but was not an overriding concern because systems were centralized and access was limited to Automated Data Processing (ADP) personnel. Security became a program of ensuring that the personnel were trustworthy and that sufficient backup of the data was located at a separate site.

As ADP systems became decentralized and the breadth of the information stored in the systems expanded, the need for a more thorough risk management program was evident. The concept of risk management was recognized as an integral part of computer security but the process of risk management lacked the sophistication of an effective tool (Glaseman, 1977, pp.105-111). The most comprehensive approach prior to 1979 was the application of checklists.

By October 1979, significant research efforts had been expended to develop the first broad-based method for evaluating a system's risks -Federal Information Processing Standard Publication (FIPS PUB) 65 of the National Bureau of Standards (NBS). The method was intended to quantify the risk analysis procedure to provide management with appropriate information to make cost effective decisions regarding safeguards (National Bureau of Standards, 1979). In the absence of any other federal guidelines, the annual loss expectancy

(or expected value) approach in FIPS PUB 65 became the de facto standard for conducting a computer security risk analysis for most federal agencies.

Although FIPS PUB 65 went a long way to establishing a quantitative approach for risk analysis, it had a number of inadequacies. First, it was oriented to large, centralized systems and did not accommodate the dispersion of data or processing. Second, the method was subjective in its application of asset valuation, threat likelihood, and occurrence impact; this introduced inconsistencies into the method. Third, the application of the procedure was disproportionately costly in terms of time and manpower. Despite these inadequacies, FIPS PUB 65 and its expected value approach to risk analysis for computer security is still used by a large number of organizations.

In January 1985, the Air Force Computer Security Program Office hosted a conference that was attended by many federal agencies that were involved in developing alternative risk analysis procedures. The topics addressed at the conference were similiar to those discussed in other non-governmental forums such as the Society for Risk Analysis, an international organization of risk analysis professionals from a variety of disciplines. One result of the conference was that the NBS and the National Computer Security Center (NCSC) agreed to assume a joint leadership role in developing and refining risk management practices. Among other decisions, NBS and NCSC established long-range plans for improving risk management practices for the federal government. These plans entailed developing and validating a framework for risk management, developing new risk management procedures, and developing metrics for the evaluation of those procedures. Each of these goals will be discussed briefly.

1. The Framework

By 1988, NBS and NCSC had organized what was to become an annual conference referred to as the Computer Security Risk Model Builders Workshop. The central theme for the conference was the modification and enhancement of a basic framework for risk management. The model was based on a concept first proposed by Stuart Katzke, Chief of the Computer Security Division of NBS, at the 1985 Air Force conference. His "strawman" included a description of the elements associated with risk management as well as the beginnings of the relationships between them. (Mayerfeld, 1989)

The goal was to refine the framework into an understandable and all-encompassing structure that would identify all the elements of the risk management process and define the functional relationships between those elements. This framework would then be used to validate a particular risk assessment model or categorize different models.

2. New Risk Management Procedures

Although the work on the framework remains incomplete, there is a compelling requirement to assess the risks to a computer system; for federal agencies, Office of Management and Budget (OMB) Circulars specifically require a periodic risk analysis be conducted. Even without federal directives, common sense dictates a carefully structured approach to examining risk exposure for federal, commercial, non-profit, and private entities alike.

Associated with an examination of an organization's risk exposure might be an enormous cost in time and effort, especially if the risk analyses were conducted manually. In essence, the cost could conceivably exceed the benefit of the results. Consequently, several

federal agencies and commercial organizations have developed automated programs to conduct risk analysis, in hopes of reducing the cost and effort required.

3. Evaluation Metrics

With the proliferation of automated products available for computer security risk analysis, the National Institute of Standards and Technology (NIST, formerly NBS) and the NCSC cooperatively sponsored the Risk Management Research Laboratory to provide facilities to examine and compare the various products. However, the evaluation and comparison of the above products is influenced by the evaluator's personal biases because of the lack of a systematic and standard comparison criteria. Lance Hoffman noted at the 1986 National Computer Security Conference:

> One significant lack today is metrics for risk analysis and risk management. There is no currently accepted set of criteria against which all methods can be compared. It is difficult to evaluate or to convey the advantages and disadvantages of a given methodology or tool when no accepted evaluation metric exists. (Hoffman, 1986, p.157)

The third long-range goal of the 1988 conference was targeted at that situation. To date, no known effort has been directed toward the development of metrics for the evaluation of risk analysis and risk management methods and the wide variety of automated tools for computer security (Pinsky, 1989).

B. THESIS PURPOSE AND APPROACH

1. Purpose

Risk analysis criteria have not been given as much attention as the risk management framework or automated tools. They are a vital component of the selection of any risk management procedure. Without the means of comparing or evaluating decision-

making methodologies, the metadecision (the decision of which method to use to make a decision) becomes arbitrary and capricious. The selection of an inappropriate method or tool could lead to excessive costs, misdirected effort, and the loss of assets. Excessive costs could result from the selection of a more expensive safeguard than is actually necessary, or from misdirected effort. The ultimate impact of an inappropriate decision could manifest itself in the loss of assets. Intuitively, the use of the wrong methodology could be worse than not performing risk management at all. Metrics could provide the means by which a framework, procedure, or tool is measured for suitability, and standardize the evaluation process of determining the most appropriate methodology for a given situation. Referring to the Katzke framework under modification and enhancement, Browne writes:

> One of the missing elements of the Framework is an exposition of the criteria that a user organization or a methodology developer needs to know in order to create, enhance, or evaluate a methodology or methodological tool.
>
> Given a purpose and an objective for risk assessment, a government agency or commercial organization should have a criteria list so that a variety of methodologies can be evaluated. For example, issues of: breadth of analysis, depth of analysis, cost of analysis, precision of results, reputability [sic], and materiality need to be explicit, and can form the basis for an evaluation of a given methodology. Different situations will call for different solutions. The trade-offs need to be known.
>
> The developers and vendors of risk management methodologies can use the same criteria to position their offerings and provide solutions that meet the needs of their client base. (Browne, 1989, pp.14-15)

The objective of this thesis is to examine existing computer security risk management methodologies, thereby categorizing their strengths and weaknesses. The ultimate purpose is to establish a standardized set of metrics that can be used to evaluate risk management methodologies and tools for their suitability to a particular organizational

situation. This objective includes several intermediate goals which will be accomplished in three phases:

- Design and develop metrics for the establishment of evaluation benchmarks for risk analysis and risk management methods.
- Apply the benchmarks developed above to compare and correlate methods.
- Apply the benchmarks to compare and correlate manual or automated risk analysis tools.

2. Approach

The need for acceptable computer security risk mana ient practices is becoming more evident throughout the federal and commercial environment because of the sophistication and complexity of today's technology and the increased value society has placed on information. Present methodologies, although suitable for some situations, have lost their appeal for several reasons.

The first and probably the most prevalent vocalized complaint is the amount of time required to complete a thorough risk analysis. When a computer system of substantial complexity is involved, teams of several people can spend months just for the risk assessment alone. Considering the entire process to be iterative, a thorough assessment could become disproportionately expensive in comparison to the cost of a periodic loss to the system.

The next problem often expressed is the lack of confidence in the probabilities available for the likelihood of threat events. This lack of confidence is derived mostly from the inherent problems of understanding probability and the methods available for producing those probabilities. For instance, it is difficult for an analyst to discriminate between the probabilities of .15 and .20 for the occurrance of an event. Also, an analyst may have a

difficult time conceptualizing a probability of 10^{-9}. Because computer technology is so young in comparison to other, more mature technologies, an established data base from which to develop historical probabilities is not available.

Another common complaint of present methods is the inability to combine qualitative and quantitative information. This particular problem permeates throughout all risk management procedures as well as throughout any particular procedure itself. For instance, an analyst can survey the stakeholders and owners of data to determine the value of that data. However, because of the nature of data, its value would most likely be expressed in qualitative terms rather than quantitative terms (such as dollars). A specific example is the value placed on classified data by federal agencies. It is expressed in terms of the level of protection required (National Computer Security Center, 1985). Although this may be suitable for the analyst's purposes initially, he eventually must combine it with the other values from an assessment. Consequently, practitioners of computer security require a methodology that resolves these problems.

Initially, our approach to this task was to compare risk management methodologies of other, more mature technologies and evaluate their techniques of dealing with the valuation of assets, determination of likelihood, and the impact significance. For instance, the insurance industry routinely places a monetary value on irreplaceable assets (such as people). Civil engineers effectively accomplish significant undertakings because they have standardized thresholds of acceptable risk. Additionally, the nuclear power industry exists, despite an extremely low threshold of acceptable risk in that industry. We had hoped to draw from the strengths of these technologies and apply their special capabilities to risk management for computer security.

Although logically, this seemed like an appropriate technique for the development of a new methodology, we immediately discovered an impasse - a lack of a technique to evaluate specific attributes of risk management methods. Without a technique to compare risk management methodologies, a framework from which we could base our research was missing. Thus, we focused our efforts on developing a paradigm that facilitates the comparison of risk management methods utilizing factors such as suitability, quality or acceptability.

Because the initial literature search failed to uncover any significant work in the area of suitability, quality, or acceptability metrics for computer security risk management, we expanded our search to a broader scope of risk management disciplines. The results of the search indicated that little work had been done in this area.

In 1978, Stephen Pollock submitted an article to the Advanced Research Institute on *'Education in Systems Science'* that explains the use of a university modeling studio to teach mathematical modeling. As part of the article, he defines and discusses twelve attributes that he uses to evaluate models produced by his students. His method is purely subjective and qualitative in nature, but the attributes could be adapted to our needs as many of methods for risk management use modeling techniques. (Pollock, 1978, pp.211-225)

In 1981, Fischhoff, et al, attempted to establish evaluation criteria for social science risk analysis methods. They were able to categorize acceptable-risk decision making methods into three groups: 1) formal analysis, 2) bootstrapping, and 3) professional judgement to which they subjectively applied seven criteria of acceptability. Again, the criteria are qualitatively defined. (Fischhoff, et al, 1981, pp.47-60)

Another effort that bears discussion for this thesis has been conducted by Merkhofer in 1985, also for the evaluation of social risk management and decision-aiding approaches. Merkhofer categorizes criteria for evaluation as either internal (which pertains to the domain of the analysis) or external (which pertains to the considerations and constraints outside of the analysis). Within this classification of the considerations for a particular decision-aiding approach, Merkhofer specifically defines criteria as "logical soundness," "completeness," "accuracy," "practicality," and "acceptability." Each of these criterion are used to assist the analyst in determining which methodology is most suitable for his given situation. Unfortunately, the interpretation of the meaning or degree of applicability for each criterion is again left to the judgement of the analyst. (Merkhofer, 1987, p.189)

Our intentions are not to imply that the judgement of the analyst is not an acceptable evaluation medium. Quite the contrary, it is the analyst that is in the best position to determine which methodology should be selected. The desire for a more objective or measurable criteria is to remove analysts' deficiencies or biases from the evaluation, thereby enhancing the evaluation of suitability of a method to a given situation. Using Merkhofer's criterion of practicality, how does one distinguish all the attributes that describe practicality for one method and compare them with those of another method using qualitative descriptions? What Merkhofer's method needs is a technique that would granulate each criterion and each attribute to a level that could be quantitatively measured.

The research conducted in the field of Software Engineering, specifically the development of software quality metrics, provides a means of quantitatively measuring quality. Software engineering metrics assist an evaluator in assessing the quality of software under development. We recognized this ability to quantitatively measure qualitative

characteristics as adaptable to the problems of risk management. The potential for this adaptation drove the literature review for the thesis well beyond the risk management and risk analysis arena to include the works of software quality pioneers such as McCall, Boehm, and others.

Utilizing their techniques, a conceptual approach for the development of our evaluation method ensued. This approach necessitates the establishment of the relationships of criteria to suitability and the definition of those criteria. Further decomposition of each criterion describes those attributes necessary to quantify a series of measurable components (McCall, et al, 1977, pp.1-2). These components form a set of metrics that are used to provide a value for each of the criteria. This thesis develops an evaluation strategy for comparing computer security risk management methodologies and tools through granulation of criteria to their measurable components.

C. THESIS STRUCTURE

In this chapter, we have established a need for an improved methodology for computer security risk management. As we conducted a survey of available methods, we became involved in the actual evaluation of those different methods, and changed our focus from developing a more effective and efficient risk management method to the task of designing and developing evaluation criteria for the suitability of risk management and risk analysis methods for computer security. Our rationale and technique as well as the scope of our literature search have been explained in this chapter.

Chapter II defines and describes our criteria, their attributes, and the relationship of the metrics. Chapter III discusses a variety of options for applying the metrics, and evaluates four

generic risk analysis methods. Extending the application of the metrics to four actual risk management tools is accomplished in Chapter IV. In conclusion, Chapter V discusses the implications of utilizing the metrics as a guide to the development of an optimal methodology, and provides concluding remarks with recommendations for further research.

II. SUITABILITY METRICS

A. ENVIRONMENT REQUIRING METRICS

In the past, risk management was performed routinely by practitioners of all disciplines based on combinations of common sense, personal experience, trial and error, and ordinary knowledge. In recent decades however, technology has driven the capabilities of this society to new and exciting achievements. With those achievements, however, technology has brought an abundance of new risks that must be understood. To understand and prepare for these risks, society demands analytical methods that allow planning, forecasting and forewarning of events which may be adverse. The scientific community has responded to this demand with numerous quantitative analytical methods for risk management and decision-making under uncertainty. (Barclay, et al, 1977) This response by itself has introduced another problem: which method is best for a particular situation?

Selection of a particular methodology is a difficult task for individuals responsible for the performance of risk management within an organization. There are several prerequisites a person must possess in order to successfully accomplish his task; the first of which is a thorough understanding of the system he is managing. This includes understanding not only the technical aspects such as hardware, software and communications, but also the non-technical aspects such as the personnel that have access to the system, the procedures they are using, and the stakeholders' concerns about that system. The second prerequisite is that he must understand the scope of the analysis he must perform and its suitability to the purposes of the organization. This entails determining how detailed the analysis should be,

how significant the treatment of uncertainty is to the result, and the appropriate form of the results for making a decision. Both of these requirements will be different for each manager, for each organization, and for each system.

The third prerequisite is to gain a thorough understanding of enough of the methods available so that the analyst can choose the one appropriate for a given situation. This requirement remains relatively constant, yet is overwhelming for one individual. Within the information resource technology discipline, several risk management methods are available for determining risks. Among these are Quantitative, Checklist, Scenario, Questionnaire methodologies, and hybrids of each. Each method has its own strengths and weaknesses depending on its application. The culmination of a risk assessment is the determination of a level of risk that is compared to an acceptable level of risk for that organization. If it exceeds the acceptable level, then a decision must be made regarding the excess risk. If it doesn't exceed the acceptable level, then no action is required. Gaining proficiency with each of these methods, as well as any new method that may be developed, is generally difficult.

Selection of a suitable methodology is further exacerbated by the complexity of large mathematical models and the amount of information required to be maintained about a system to complete the analysis (Barclay, et al, 1977). As a result, the mechanics of performing risk analysis procedures has evolved from manual paper and pencil drills to the use of the computer for storing, computing and analyzing the entire process. Federal agencies and commercial vendors have responded to the need for risk management packages with a multitude of products from simple database storage to sophisticated analysis packages employing artificial intelligence and expert systems.

The Risk Management Research Laboratory, sponsored by NIST and the NCSC, provides facilities to examine and compare various commercially available risk management products. The lab has become a central repository of risk management tools which facilitates the accessibility of the products for evaluation and comparison by federal agencies. During the course of our research, we had the opportunity to visit the lab and examine over two dozen risk management packages. Although the lab provides the convenience of having these products at one location, we experienced the frustration and confusion of the absence of an evaluation and comparison criteria.

The lack of a systematic and standard comparison criteria makes any attempt at evaluating tools arbitrary and indiscriminate. It is apparent that most managers in the position of selecting an automated risk management tool (or more generally, selecting a methodology for risk management) have little guidance or rationale to make a suitability determination. If the NIST/NCSC lab could incorporate a technique which provided a standardized evaluation method, its functionality would be greatly enhanced.

In 1982, Congress passed the Risk Analysis Research and Demonstration Act which was "...intended to encourage researchers to 'define criteria and standards to guide the development and use of risk analysis' and 'to improve the methodologies' used in risk analysis." (Shrader-Frechette, 1985, p.5) As a result of our literature search, it became apparent that significant efforts had gone into the development of improved risk management methodologies, although little effort has been devoted to the definition of criteria and standards.

The application of our criteria for evaluation of computer security risk management methodologies follows those of Merkhofer (1987), but diverges when we reduce the

subjectivity of the criteria by the introduction of metrics. We have examined many decision-making methodologies in both the risk management and decision analysis disciplines. During our examination, categories of characteristics were identified that were pertinent or important to a specific situation, and we defined several qualitative attributes which describe suitability.

What is suitability? We have defined suitability as those characteristics of a risk management methodology or tool that are pertinent and appropriate for the requirements of a particular person, organization, system, and/or situation. Use of predefined criteria for evaluation of tools and methods allows the user to make comparisons based on a standard continuum and to tailor an evaluation to his particular needs. As it will become apparent, the criteria and their attributes are qualitative in nature thus they submit themselves to subjective interpretation. At this level, the evaluation criteria for suitability are just as difficult to discern as suitability in totality. We remedied this by further decomposition of each attribute, into what are referred to as metrics, which provide a quantitative description of the suitability criteria. Each of these criteria are discussed in detail with their attributes and metrics in the next section.

B. EVALUATION METRICS

The intent of this thesis is to provide a means of quantitatively measuring the suitability of a method or tool for a particular organization for the purpose of evaluating methods or tools. The procedure that will be used to accomplish this is listed in Table 1 below:

TABLE 1. Steps for Measuring Suitability

1. Establish a set of criteria that describe a method's suitability.
2. Define the suitability criteria in terms of related attributes.
3. Specify metrics that describe the presence of the attributes.
4. Make a quantitative statement of the appearance of the suitability criteria by determining the ratio of actual occurrences of a metric to the number of possible occurrences.
5. Use the derived quantitative values for each of the criteria to evaluate and compare the variety of methods and tools available to the organization.

By taking this approach, the analysis of suitability becomes standardized, flexible, and expandable. Standardization of the evaluation method will provide continuity in comparative figures. For instance, adaptability is judged consistently across all methods or tools and can therefore be used as a comparative measure. The analysis becomes flexible because the organization making the evaluation can choose which of the criteria best define suitability for their organization. As an illustration, in an organization where there are a wide variety of computer system configurations, adaptability and consistency might be more important than useability. In this case, greater weight could be added to adaptability and consistency. As the definitions of suitability are refined, the analytical method presented is expandable by simply adding metrics, attributes, or criteria as they are developed. This chapter will detail the procedures for achieving the steps one through three from Table 1 while Chapter Three will demonstrate the application of steps four and five.

To maintain consistency throughout the thesis, a number of definitions are provided:

- Criterion. A characteristic or trait which indicates a level of suitability for a risk management method or tool. Criteria actively contribute to the suitability of a method by their presence.

- Attribute. A quality which, when taken with other attributes, describes, defines, or judges a criterion.

- Metric. A boolean question which provides a specific measure for each of the attributes related to the criteria. Most of the metrics are subjectively determined and are intended to provide a direct evaluation of the degree of existence of an attribute. The unit of measure for the metrics is expressed as a ratio of the number of occurrences of the metric to the number of possible occurrences.

Following the procedures listed earlier, our first task was to specify the criteria that describe suitability. Adjectives that typify suitability were listed and then grouped according to their relationships with each other. Each of the groupings constituted a criterion while the adjectives in the group became the attributes of the criterion. Finally, we defined the attributes in terms of a comprehensive set of rudimentary, boolean questions. These questions comprise our metrics. For example, the question, "Is there a standardized interface?" is a metric of the attribute 'ease of use.' The questions were worded to imply a positive aspect of the attribute if answered in the affirmative. In this way, the existence of that metric implies that the characteristic contributes to the presence of the attribute.

The result of this process was seven criteria composed of between two and four attributes. The criteria are: *consistency*, *useability*, *adaptability*, *feasibility*, *completeness*, *validity*, and *credibility* (see Table 2). Each will be discussed in detail in the paragraphs below.

The relationship between criteria, their attributes, and metrics described above supports formulation of a simple mathematical relationship between the metrics and their associated

TABLE 2. Suitability Criteria

Consistency. Given a particular system configuration, results obtained from independent analysis will not significantly differ.
Useability. The effort necessary to learn, operate, prepare input, and interpret output is generally worth the results obtained.
Adaptability. The structure of the method or tool can be applied to a variety of computer system configurations (and the inputs can be easily updated as they periodically change).
Feasibility. The required data is available and can be economically gathered.
Completeness. Consideration of all relevant relationships and elements of risk management is given.
Validity. The results of the process represent the real phenomenon.
Credibility. The output is believable and has merit.

criterion. When an evaluation of a method or tool takes place, the resultant measurements can be viewed as a set:

$$(m_1, m_2, m_3, \ldots, m_n) \tag{1}$$

Each element, m_i, represents a boolean measure of the presence of that aspect of the attribute in the criterion. To develop a mathematical expression of each attribute, A_j, held by a criterion, the positive boolean metric values are summed and expressed as a ratio:

$$A_j = \frac{\sum_{i=1}^{n} m_i}{n} \tag{2}$$

For example, if six of the nine possible metrics of the attribute *scope* are positive, the value of A_j is calculated by dividing the number of positive metrics (6) by the total possible number of metrics (9) for that attribute (.666).

A criterion, C_k, in turn, is expressed mathematically as a ratio of the sum of the attributes' values to the number of attributes:

$$C_k = \frac{\sum_{j=1}^{p} A_j}{p} \tag{3}$$

With expressions of criteria described, a description of the suitability, S, of a method or tool can be expressed as an ordered set of the values of its criteria:

$$S = (C_1, C_2, C_3, ..., C_k) \tag{4}$$

Finally, an evaluation of a method is simplified by comparison of the suitability expressions of various risk management methods or tools. For instance, the suitability of a particular method can be compared to another method by examining the elements of the two vectors, S_1 and S_2. A 'suitability index,' S_t, can be derived by attaching appropriate weights, W_q, to the criteria and summing the results for all criteria. The suitability index is then expressed as a ratio:

$$S_t = \frac{\sum_{q=1}^{r} W_q C_q}{r} \tag{5}$$

The approach described above is the most simple application of the evaluation technique. Additional methods of application of the evaluation metrics are described in Chapters III and IV.

Each of the seven attributes will now be described in turn:

1. Consistency

The concept of consistency, as applied to suitability of risk management methods or tools, implies an ability to duplicate the results of the process. In other words, given a particular system configuration, results obtained from independent analysis will not significantly differ. Consistency was determined to have the attributes of *reliability* and *consistent terminology* (see Figure 1).

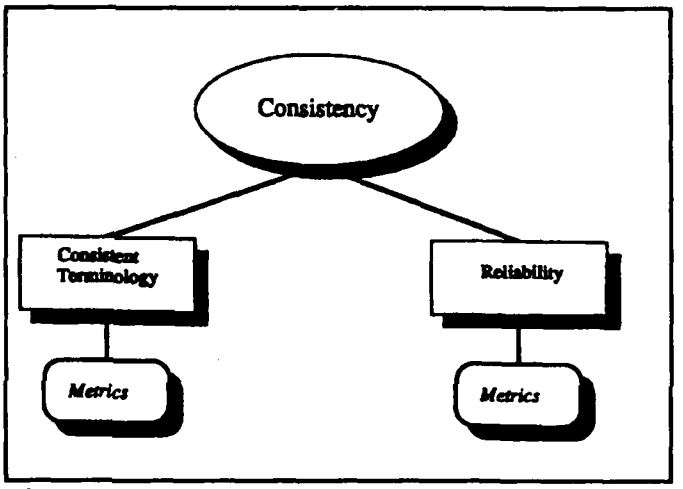

Figure 1. Consistency

A key component of reliability is objectivity, or the reduction of subjectivity in the process. Objectivity reduces the wide amount of variance that could occur as a result of personal biases. A variety of researchers including Kahneman (1982) and Gardner (1989) have provided writings on the subject of bias and its effect on decision-making.

> When laypeople are asked to evaluate risks, they seldom have statistical evidence on hand. In most cases, they must make inferences based on what they remember hearing or observing about the risk in question. Psychological research ... has identified a number of general inferential rules that people seem to use in such situations. These judgmental rules, known as heuristics, are employed to reduce difficult mental tasks to simpler ones. Although they are valid in some circumstances, in others they lead to large and persistent biases with serious implications for decision making.... (Slovic, et al, 1980, pp.464-465)

Fischhoff (1982) describes a number of methods of reducing or eliminating biases from the decision making process. The more the process reduces biases in the analysis, the more consistent the results will be between analysis teams.

Another aspect of reliability is the reduction of uncertainty. Uncertainty, by definition, is a significant part of risk and the risk management process. For results of the process to be consistent across analysis teams, an effective mechanism must exist to reduce the impact of uncertainty.

To control differences in interpretation of what the process is asking for and what the product represents, a uniform set of terminology must be established between the analyst and the process. This involves providing understandable definitions to the analyst which will be consistently applied throughout. Terms should be used in their commonly accepted context and unambiguously defined to avoid misinterpretation.

Table 3 below and Template 1 in Appendix A describe the criterion of consistency, list its attributes and their associated metrics, and represent the mathematical relationships between them.

TABLE 3. Consistency

ATTRIBUTES	METRIC	YES	NO	VALUE
Reliability	1. Does the process provide a mechanism to reduce the introduction of personal bias?			
	2. Does the method provide a mechanism that reduces the impact of uncertainty?			
	RELIABILITY Attribute Value $(A_1)=(m_1+m_2)/2$			
Consistent Terminology	1. Is a standard language established?			
	2. Are the method's elements defined for the user?			
	3. Does the method request input in designated units?			
	4. Is the input requested unambiguous?			
	CONSISTENT TERMINOLOGY Attribute Value $(A_2)=(m_1+m_2+m_3+m_4)/4$			
Metric Value	CONSISTENCY Metric Value ($(A_1+A_2)/2$)			

2. Useability

Useability is defined as the value of the effort necessary to learn, prepare input, execute the process, and interpret output. The four steps described above represent the interface between the analyst and the process. Hence, for a process to be useable to an operator, it must be *understandable*, *easy to use*, *simple*, and effective at *handling errors* (see Figure 2).

Understandable refers to the ability to comprehend the underlying premise that supports the method. The need for understandability is underscored by Fischhoff, et al:

> An approach should not make matters worse by obscuring its internal functioning. All those whose fate it may affect have a right to ask: What are its underlying assumptions? What are its political and philosophical roots? ... What inputs were used? What computational procedures were followed? How much uncertainty surrounds the entire enterprise? To be valid an approach must provide answers to these questions.... The unexamined approach is hardly worth using. An approach that fails to test its effectiveness and clarify its prejudices is not to be trusted. (Fischhoff, et al, 1981, pp.56-57)

The user of the method does not have to grasp all aspects of the process, but must have an appreciation for what will be asked of him and how the results of the process relate to the inputs. Can he understand what decisions are expected of him, and what

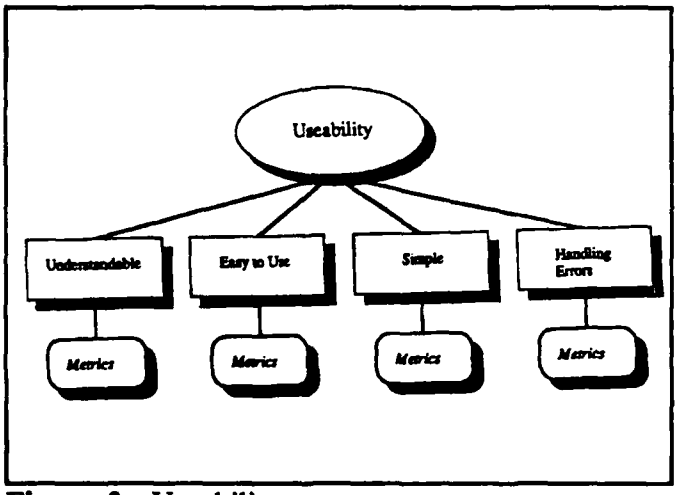

Figure 2. Useability

decisions are being made by the process? Can he follow the process (in general terms) to appreciate the rationale of the output?

The next attribute that contributes to the definition of useability is ease of use. Although the definition of ease of use could be considered intuitive to the reader, our definition embraces a process that is well structured and logically sequential. The result is a procedurally unconvoluted process. Along with this structured process, there must be a consistent interface that allows the analyst to concentrate on his task rather than on the process itself.

At the risk of appearing trivial, we differentiate ease of use from simplicity. The distinction between ease of use and simplicity is demonstrated by the modern telephone system. The underlying technology of the telephone is far from simple with its elaborate networks, relays, and switches. Yet, the telephone is easy to use because of its consistent interface and the concealment of this technological complexity from the user. Simplicity implies that the complexity of the process is concealed without obscuring the premise. No special training should be required of a simple system, nor should an expert knowledge base be required of the user to operate it.

The last attribute for useability is error handling. Previous attributes of useability dealt with all three phases of a procedure; input, processing, and output. Error handling, however, is focused on input only. The method or tool becomes more useable when input errors are readily identified and the resolution of the errors are facilitated. Another aspect of error handling is the sensitivity of the process to insignificant data accuracy errors. This sensitivity must be accommodating for those errors that don't

impact the final output. Many decision analysis methods, for instance, use orders of magnitude to alleviate the concern over minor accuracy errors. Cooper, in describing an Annual Loss Expectancy method for risk management, provides a table of threat frequency ranging from 100 times per day to once every 300 years (stepping by a factor of 10). His reason for utilizing orders of magnitude is to "...operate within the bounds of the input accuracy and to minimize computational complexity and time expended." (Cooper, 1989, p.28)

Template 2 in Appendix A describes the criterion of useability, lists its attributes and their associated metrics, and represents the mathematical relationships between them.

3. Adaptability

Adaptability describes suitability in so much as the structure of the method or tool can be applied to a variety of computer system configurations. Additionally, stored data can be easily updated as it periodically changes. The attributes that describe adaptability are *portability* and *modifiability* (see Figure 3).

Portability is the ability to apply the process across a variety of computer system classifications and configurations. For example, a method that is designed for a single site mainframe environment and does not pertain to a highly distributed environment would receive a poor ranking in portability. Within a single system, portability also applies to the mode by which the system processes jobs; i.e. batch versus interactive processing. Another facet of portability is its applicability to not only the system but also to the immediate environment of the subject system. Does changing the

location of the computer system cause the method to become ineffectual or will it still apply because it considers the system's environment? And finally, system life cycles must be taken into consideration to reflect the changing disposition of all systems as they progress from inception to replacement.

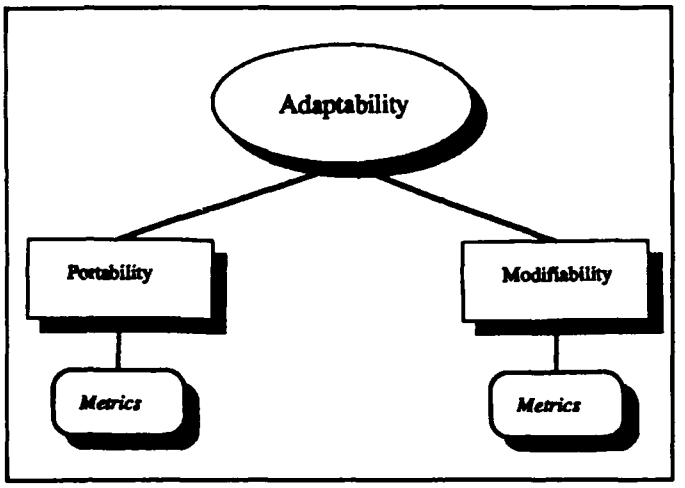

Figure 3. Adaptability

Just as it is important for a method to be portable because systems change, a method should also be capable of retaining original input values for recall and segmenting calculations into logical partitions. This capability would allow for examining the effects of a particular change while holding most other variables constant - a 'What if?' capability. Katzke, in his summary of issues addressed by the 1985 Risk Analysis Workshop, listed a set of desirable properties of a risk management method. These properties included automation, which "... maintains [sic] database of system descriptions...." and a capability which "... allows iterative safeguard cost benefit analysis." (Katzke, 1985, p.15) We have identified this capability as modifiability which addresses that portion of adaptability that assists the analyst in examining alternatives or options.

Template 3 in Appendix A describes the criterion of adaptability, lists its attributes and their associated metrics, and represents the mathematical relationships between them.

4. Feasibility

The criterion of feasibility characterizes the input data to a method. It is concerned with the amount of data that must be collected, the economy of gathering that data, and whether the data is obtainable without extraordinary measures. The attributes that describe feasibility are *availability*, *practicality* and *scope* (see Figure 4).

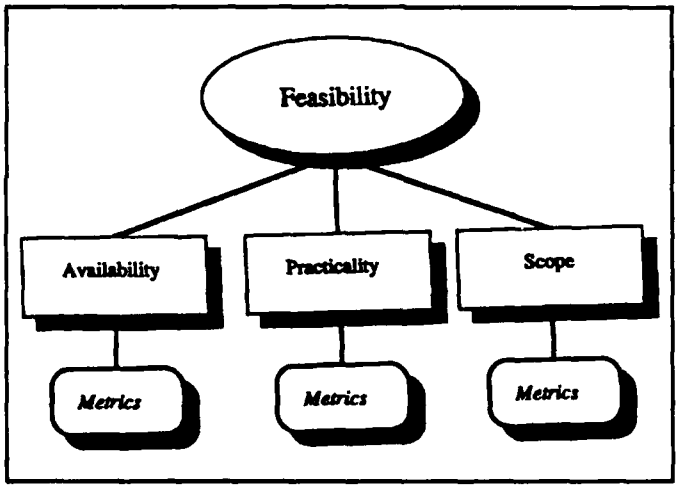

Figure 4. Feasibility

Availability distinguishes that data which is accessible from within the organization (either from expert knowledge or historical records) from that data which must be collected externally. This is not to imply, however, that all data requirements are fulfilled from internal sources for it to be available. Data required which is not organic should be convenient to collect, thereby conforming with the intent of the availability attribute.

Practicality, on the other hand, is concerned with the economics of gathering the required data. Data collection for any system is not free. The cost of performing the

data collection must be estimated prior to initiation and a conscious decision to expend the necessary effort and time to accomplish this task has to be made.

> Practicality means that the analysis can be conducted in the real-world, problem-solving environment using available resources and information.... Practicality is obviously influenced by the availability of required inputs and the extent to which the approach is flexible in its ability to use different types of input data. (Merkhofer, 1987, p.191)

Practicality is also the required precision of the data from the method's perspective. As a negative example, if a method demands evaluation of asset value to the nearest cent, it might be considered impractical for a situation where assets are valued in the millions of dollars.

For a method or procedure to be feasible, it should also allow the analyst to select or determine his own level of detail or precision. This is what we refer to as the scope of the analysis. The analysis of a system is shaped by the breadth that the process allows the analyst to select. Scope is the means by which the bounds of an analysis are defined, thereby influencing the acceptability and usefulness of a method. Stevens and Weiner noted the difficulty experienced by Department of Energy (DOE) officials with their risk assessments, "Most of the DOE professionals surveyed indicated that they had difficulty in determining the scope of a risk assessment..." (Stevens and Weiner, 1989, p.475). Because of scope's significance, methods should consider the different segments of a system such as hardware, software, data, personnel, procedures, communications, and the environment that the system functions within. Scope has an impact on other criteria as well, and will be further described below.

Template 4 in Appendix A describes the criterion of feasibility, lists its attributes and their associated metrics, and represents the mathematical relationships between them.

5. Completeness

Completeness is defined as providing comprehensive coverage of all considerations of the risk management problem. To be complete, a method should regard all components and the attributes of those components of the system. Additionally, the process should examine the computer system at the level of detail desired by the analyst. The attributes that describe completeness are *scope*, *elements* and *element attributes* (see Figure 5).

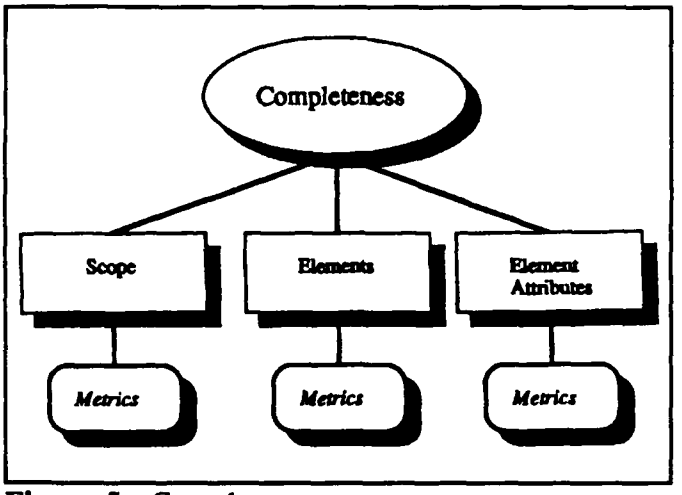

Figure 5. Completeness

For completeness to be achieved, a comprehensive examination of the key elements of risk management should be given. To date, a consensus has not been reached by practitioners in the computer security risk management discipline to define an exhaustive list of the key elements and their attributes. Probably the best effort to list and define those elements and attributes is provided by Mayerfeld. In his paper, he prepared a composite of elements and attributes developed by four independent working groups of the 1988 Computer Security Risk Management Model Builder's Workshop:

> Within a scoped system three central elements operate to determine the risk of the system. These primary elements are assets, threats, (threat agent/threat event), and vulnerabilities/safeguards. An asset is defined in terms of its value, a threat event in terms of its undesirability, and a safeguard in terms of its effectiveness.
>
> ... The asset is the entity whose risk we measure (normally in aggregate sets of assets), and whose protection we desire. The critical attribute of an asset is value....
>
> Undesirable outcomes that can befall assets are initiated or accomplished by or through threats, that can change an existing state to a new, potentially less desirable, state. The term 'threat,' ... refers to both threat agents and threat events.... The defining characteristic of a threat agent is its potency, which is an aggregate measure of its potential to instigate a threat event. The critical attribute of a threat event is its undesirability. A complete analysis of the outcomes and consequences of a particular threat event is necessary in order to fully assess how undesirable it is....
>
> Protection of assets is accomplished by the existence or implementation of safeguards, or countermeasures. Safeguards can be active or passive, and are often defined as part of the scoped system. The critical attribute of a safeguard is its effectiveness. Any measure of effectiveness also requires an analysis of the threat events against which the safeguard is said to be effective.
>
> The interactions between the three primary elements determine the outcomes or consequences, that are likely to result, and their likelihood. The key attribute of an outcome is its severity." (Mayerfeld, 1989, pp.10-11)

Using Mayerfeld's definitions, the elements that should be considered by a complete method are assets, threat agents, threat events, safeguards, vulnerabilities, and outcomes.

The ability of the user to control the level of detail of analysis and consider all aspects of the system, defined earlier as scope, is an important aspect of completeness. To satisfy the needs of the organization, the process must be able to adapt to the needs of the analyst to examine his system in the detail that he desires and analyze any and all aspects of the system.

Template 5 in Appendix A describes the criterion of completeness, lists its attributes and their associated metrics, and represents the mathematical relationships between them.

6. Validity

Validity is measured as the extent to which the results of the process represent the real phenomenon. In other words, the conclusion of the analysis resembles what is experienced in reality. For instance, if the process recommends a perimeter fence for a computer center located on the twelfth floor of an office building, the process is not valid. The attributes that describe validity are *relevancy*, *scope* and *practicality* (see Figure 6).

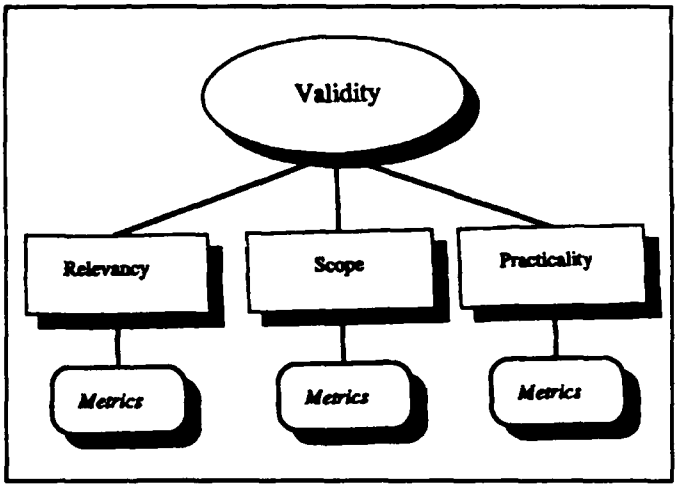

Figure 6. Validity

Relevancy means that the results are meaningful to the system. To avoid irrelevant conclusions such as the example cited above, a process should provide categories of solutions rather than specific recommendations. A recommendation of 'access control' (as a category) rather than a specific recommendation of 'a fence' would be less likely to be irrelevant, although not necessarily preferrable. Results from a method should relate to significant areas of need and incorporate mandated security requirements to maintain their relevancy.

To be valid, the method must be able to provide the scope, as defined in paragraphs above, at the level of detail required by the analyst. Scope determines the extent of the detail used by the process and is necessary. Otherwise, the method's perspective does not reflect the intent of the analyst.

Practicality is conversely related to validity. The lack of practicality suggests that biased or erroneous input data would be provided in the place of data that could not be economically gathered. Presumably, the result in this case would be less valid than if data collection was practical.

Template 6 in Appendix A describes the criterion of validity, lists its attributes and their associated metrics, and represents the mathematical relationships between them.

7. Credibility

The credibility of a particular method or process has significant bearing on the acceptability of its conclusions. If any aspect of the process is questionable, the entire process will be suspect to the users. The attributes that describe credibility are *intuitiveness* and *reliability* (see Figure 7).

The input, process, and output of a method or tool must have a natural feel that will instill and maintain the confidence of the analyst. This intuitiveness is measured in terms of the amount of information available to the user. If the method delineates the relationships between elements, the analyst will be able to understand how values are derived and is likely to find them plausible. In the same way, the output of the process must have a perceptible relationship to the data that was provided. Finally, if the method fails to analyze an aspect of the system that the user knows to be flawed

or require protection, his confidence in the method will be diminished.

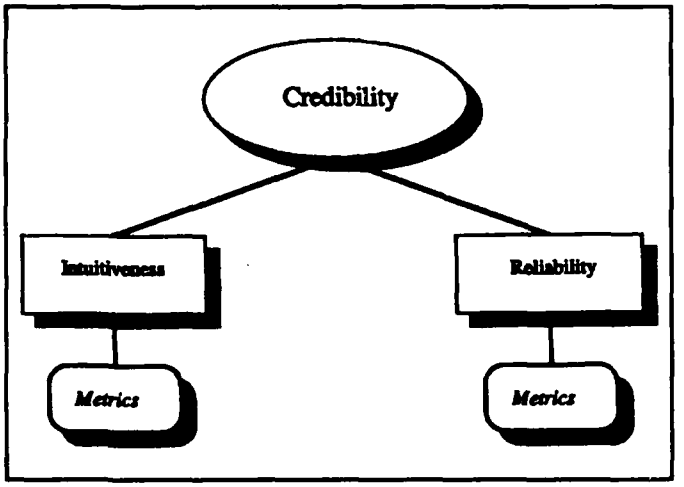

Figure 7. Credibility

Reliability as previously described for consistency also pertains to credibility. If different results are returned with the same data on different occasions, the method will hold little plausibility to those interpreting the output. The reliability of the method, that is, the ability to obtain repeatable results, has a direct bearing on the credibility of a process.

Template 7 in Appendix A describes the criterion of credibility, lists its attributes and their associated metrics, and represents the mathematical relationships between them.

C. ANTICIPATED RESULTS

As described, it is apparent that not all of the criteria can be maximized simultaneously. This is because some criteria are maximized at the expense of others. For example, in order to achieve useability, some degree of completeness will likely be sacrificed. Determining the *best* risk management tool or method will require trading one desirable trait for another. Merkhofer observed the same phenomenon:

> Notice that the various evaluation criteria are not independent. Weaknesses in some areas (e.g. logical soundness) are likely to preclude strengths in others (e.g.

> acceptability). Furthermore, some of the goals may be in conflict. It may be easier to derive a logically sound decision rule by leaving out certain awkward issues and thereby sacrificing completeness. If no approach does, or even can, satisfy all of the criteria and if their respective strengths and weaknesses lie in different areas, then the choice of an approach will require tradeoffs. (Merkhofer, 1986, p.192)

Consequently, it is felt by the authors that a categorical statement of which method is the "best" cannot be made. Suitability of a method can only be determined by careful consideration of the needs of the organization with respect to the attributes and criteria described in this chapter. After that consideration, a choice of which tool is most suitable can be made. Chapter III will offer alternative procedures for applying the metrics and demonstrate a simple approach to the comparison of selected generic risk management methodologies.

III. METRICS APPLICATION

A. INTRODUCTION

Chapter II established a means by which the difficulties of evaluating the suitability of risk management methodologies can be managed through the utilization of metrics. This chapter will augment the metric method of Chapter II by proposing alternative application options, suggesting a procedure for their use, and finally, applying the metrics to four underlying risk analysis methods. The results are examined with regard to the usefulness of the metrics by comparing them with intuitive predictions. We found a strong correlation between the predictions and the observations drawn from the results of the evaluation. This in turn provides our rationale for further application. Through this test of our evaluation method, we ascertained that the application of the metrics had a logical basis and provided reasonably accurate results.

B. APPLICATION ALTERNATIVES

There are a multitude of alternatives available to the analyst for the application of the metrics in an evaluation. The choice of which alternative to use depends upon the experience of the analyst and his requirement to tailor the evaluation to his organizational needs. Three potential linear models are discussed below.

1. Normative Relationship Model

The most rudimentary approach to the evaluation process uses a linear model with normative relationships between criteria and between attributes. The metrics are used as a boolean measure of the presence of an attribute. The values of the criteria are based only on the attribute quotients. The coefficients of criteria and attributes have a value of one, thus providing a uniform representation of suitability. The application of this model was described mathematically in Chapter II (formulas (1) through (4)).

There are a number of advantages to the use of this method, the most significant of which is its simplicity. In describing the process of modeling a phenomenon, Wan advocates that simplicity contributes to productivity in the initial stages of development:

> For any particular phenomenon, we begin by identifying a few important questions of interest and the main factors influencing the answers. We may have more contributing factors; but it is usually more productive to keep things simple initially and consider other issues after success with the simpler problem. (Wan, 1989, p.1)

As such, an analyst can utilize the normative relationship model without knowledge or concerns about the relationship between criteria. With all criteria and attributes considered equally, the distinction between strengths and weaknesses is manifested in the numerical ratings without complicated analysis. Additionally, without subjective assessments of the prominence of each criteria, minimization of bias is assured. Nevertheless, there may exist a desire or requirement to weight one criteria over another.

Application of the normative model is demonstrated in Chapters III and IV.

2. Significance Coefficients Model

In the application of this approach, the metrics remain boolean statements of the presence of an attribute, however, a significance coefficient, W, can be applied to either the attribute or criteria as determined by the need and desire of the analyst. The significance coefficient is established by the analyst to meet organizational needs with regard to risk management. For instance, if an organization has no professional computer security staff, and risk management will be conducted by personnel unfamiliar with those practices, useability might be weighted more heavily than consistency. Closer scrutiny may suggest that, 'ease of use' and 'understandability' may be more beneficial to this organization than 'portability' or 'modifiability'.

When attributes are weighted, metric values are collected as noted in Chapter II. The metrics remain boolean in nature and represent the presence of a particular attribute. Following the measurement of each metric, attributes, A_j, are described as in formula (2) (repeated below) where the attribute is the sum of its positive metrics divided by the total number of metrics:

$$A_j = \frac{\sum_{i=1}^{n} m_i}{n}$$

After determining values for each attribute, the significance coefficient, W_j, is attached to each attribute. The coefficient is established by the analyst relative to his organization's needs as discussed earlier. A criterion, C'_k, is represented by the set of its attributes:

$$C'_k = (W_1A_1,\ W_2A_2,\ W_3A_3,\ ...,\ W_nA_j) \tag{6}$$

The metric value of the individual criterion is mathematically expressed as:

$$C'_k = \frac{\sum_{j=1}^{n} W_jA_j}{n} \tag{7}$$

Finally, as presented in Chapter II (formula (4) duplicated below), suitability, *S*, becomes an expression of the set of values for the criteria from which an evaluation can be made:

$$S = (C_1,\ C_2,\ C_3,\ ...,\ C_k)$$

If the analyst requires only a more general evaluation of the risk analysis tools available, the evaluation criteria rather than attributes could be weighted. If this is the case, formulas (1) through (3) apply, where criteria are expressed as a simple ratio of its attribute's value to the number of attributes as in formula (3) below:

$$C_k = \frac{\sum_{j=1}^{n} A_j}{n}$$

After attribute values are determined, a weighting value, W_k, is assigned to each attribute. The suitability, *S'*, of a tool is then represented by the set:

$$S' = (W_1C_1, W_2C_2, W_3C_3, ..., W_kC_k) \quad (8)$$

3. Metric Utility Model

In a situation where two tools are closely matched in a particular category, one tool can be distinguished from another more readily by employing utility techniques with the metrics. Such techniques or utility functions may employ various ranges of acceptability, for instance, one to five represents the worse to the best response respectively. Utility is defined as the decision-maker's measure of the value of a particular alternative, taking into account his preference for return as opposed to avoiding risk. Von Neumann and Morgenstern developed an approach to decision making in which the decision-maker attempts to choose the alternative that will maximize his expected utility (Von Neumann, 1947, pp.15-31).

Another technique could utilize fuzzy set descriptors to characterize the uncertainty involved in assigning values to metrics such as 'Always', 'Usually', 'Never', etcetera. Zadeh describes the practicality of fuzzy techniques:

> In recent years, however, it has become increasingly clear that uncertainty is a multifaceted concept in which some of the important facets do not lend themselves to probability-based methods. One such facet is that of fuzzy imprecision, which is associated with the use of fuzzy predicates exemplified by small, large, fast, near, likely, etc. The question is: How can one express the meaning of this proposition through the use of probability-based methods?... It is questions such as these that motivated the development of fuzzy-set-theoretic techniques for dealing with problems in which uncertainty derives from fuzzy imprecision.... The techniques for dealing with information which is both fuzzily imprecise and probabilistically uncertain have a direct bearing on the important problem of inference from commonsense knowledge and its application to decision analysis. (Zadeh, 1988, Foreword)

An example of the application of fuzzy methods to computer security risk analysis can be found in the works of Bruce, Kandel, and Avni (Bruce, 1988, pp.17-49) and in the works of Schmucker (Schmucker, 1984).

In a case where weighted metrics are employed, the coefficient, W_i, is assigned based on the analyst's selection of a utility function. Assignment of the weight is made at the time the metric is evaluated and is used in the calculation of the value for each attribute A'_j:

$$A'_j = \frac{\sum_{i=1}^{n} W_i m_i}{n} \tag{8}$$

Attributes, once computed, are used as described in formulas (3) and (4) for comparative purposes. This model can be used in combination with the normative relationship model or the significance coefficient model to further accommodate known idiosyncrasies within a particular organization.

C. ANALYSIS PROCEDURE EMPLOYED

The greatest benefit derived from any of the weighted approaches is the flexibility they allow an analyst or organization. They provide a means of distinguishing the criteria that are more significant to a particular requirement. Although the use of weighted procedures for the evaluation in this chapter was a viable option, a fictitious corporate situation would have to be created that prejudged specific criteria. We felt that this was beyond the scope of this thesis and would have clouded the issue of comparison.

Consequently, the unweighted normative relationship model was selected because it would not encumber the analysis.

To reduce the volume of paper and forms associated with the suitability evaluations that were conducted for this chapter and to facilitate the comparison of the responses to the metrics, Templates 1 through 4 in Appendix B were utilized. They provided a single, adjoining view of all of the metrics, attributes, and criteria values for each method. Through the use of these forms, a quick comparison of four different risk management approaches or tools could be easily accomplished. Template 5 in Appendix B provides a comprehensive view of the criteria and attribute values for each method evaluated.

The mechanics of completing the templates is quite simple. If the answer to a metric is determined to be positive, the appropriate box is marked. Blanks indicate a negative response to the metric. Upon completion of all the metrics for a single method or tool, the values are calculated by summing the marked metrics for each attribute and dividing the summation by the total number of metrics for that attribute as in Formula (2). Likewise, the value for the criteria is calculated the same way. The result is placed in the corresponding box. When each method or tool has been subjected to the metrics, the evaluation is concluded by comparing the results horizontally for each criterion or attribute as desired. This is the technique implemented in the next section.

D. METHODOLOGY EVALUATION

1. Selection Rationale

Risk management is meant to include the complete decision process from risk analysis to the decision of safeguard selection. Although the metrics have been designed to measure risk management tools, they can also be used to evaluate a major segment of the risk management process, risk analysis. Underlying risk analysis methods are distinct and recognizable, while risk management (which encompasses risk analysis) includes a host of decision-making process choices. The multitude of choices obfuscates a general classification of risk management techniques. Therefore, we make the general distinction between risk management methods by the risk analysis technique contained therein. For this reason we chose to conduct our evaluation on risk analysis techniques, making the assumption that it was accompanied by a generic decision-making process.

A further description of each of the methods will assist the reader in adjusting to the same level of perception of the methods that is used for the predictions. To conduct a test of the metrics' utility, we selected four simple, intuitively understandable methods of conducting risk analysis. Those descriptions and the predictions are provided in the paragraphs below.

2. General Description of Methods

a. Quantitative Methods

There are a variety of distinctive quantitative methods for risk management. To apply our evaluation techniques to each of these would not significantly

enhance the comparison of the underlying methods. Consequently, we have chosen to represent this category by the most commonly encountered implementation of quantitative methods, Annual Loss Expectancy (ALE).

ALE is a method of characterizing the foreseeable losses that a system might incur through the use of statistical probability. Each asset in a system is individually considered with its associated threats. Frequencies (f) of occurrence are estimated along with the potential cost (c) or value of the loss. The anticipated annual loss for that asset from that specific threat is derived by taking the product of the frequency and the associated cost. To derive a comprehensive view of system vulnerabilities under the ALE method, all assets and threats must be considered, and their associated loss computed for each combination of assets and threats. The resultant figures are compared and ranked from highest expected loss to lowest to readily identify where the greatest potential losses could occur. (Hutt, 1988, pp.24-25)

The product of the ALE method must be augmented with additional decision-making techniques to complete the risk management process. The risks that the system face can be ranked according to organizational goals, and safeguards can be selected to provide appropriate protection. This selection is usually predicated on a minimum cost or maximum benefit evaluation of the ALE.

Nearly every application of ALE uncovered during our research provided a simplified method for handling uncertainty. Uncertainty, or the lack of empirical information about frequencies of the occurrence of threats and the actual cost of some segments of a system, was handled through tables which defined frequencies or cost in

terms of orders of magnitude. By selecting from an order of magnitude table, the user need only be accurate by a factor of 10 (or any other order of magnitude applied to the table).

b. Checklist

A technique that was popularized during the 1960's and early 1970's, checklists approach the process by attempting to anticipate insecurities in the system and checking for safeguards to ensure protection. Consisting of a series of questions intended to test the existence of a safeguard, a series of "no" answers may indicate a trend in risk to the system that requires additional protection. Checklists provide a means of identifying weaknesses in a system, but lack a means of decision-making for the security manager.

Because checklists are, by their nature, predefined, they tend to be inflexible in their ability to accommodate changes to the system. They are usually designed to be used in a broad variety of system installations (e.g. centralized batch process, distributed terminals, LAN connected PC's, etc.), which makes them more difficult to adapt to one particular system configuration than other methods such as ALE or scenario. Krauss attempts to reduce the inherent inflexibility of his checklist by allowing the user to add and remove pages of the checklist as required:

> The SAFE approach provides something between an inflexible formula and an expensive, completely custom-made evaluation procedure. SAFE represents a feasible point between these extremes and, because a reasonably high degree of custom-tailoring is permitted, the approach enables the vast majority of installations to go a long way toward meeting many important security requirements. (Krauss, 1972, p.5)

c. Scenario

The scenario methodology is suggestive of a 'think tank' technique. A group of stakeholders combine ideas, knowledge, expertise, and intuition for the purpose of analyzing, developing, and preparing possible alternatives to security related situations or events. The entire process is free form - that is, it may be structured and strongly guided toward the threat event it examines, or it may be loosely bound to a general vulnerability, safeguard, or combination thereof that the analysis team desires. Cooper describes scenarios as a "brainstorm" atmosphere that fosters a synergistic effect for identifying vulnerabilities which might otherwise go undetected (Cooper, 1989, pp.36-37).

The development of a scenario analysis is initiated by a selective team of personnel involved with a particular system in one fashion or another. Ideally, each member would be a specialist from each aspect of the system. For instance, a typical team for a centralized mainframe computer center might consist of the operations manager, maintenance representative (hardware and software), operating system specialist, communications specialist, programming or development representative, and the security manager. Each member contributes a slightly different perspective to the discussions.

As the discussions progress, a situation is developed that targets a particular asset or component of their system against a specific vulnerability or a combination of vulnerabilities. The case is developed along several possibilities to determine the most likely impact should the event actually occur. Additionally, an examination of safeguards that would protect against those vulnerabilities or would lessen the overall impact is usually included. The goal of the scenario team might range from

the simple discovery and analysis of a vulnerability to an extensive decision-making process for the selection of a safeguard. (Quade, 1975, p.118)

We considered the scenario methodology to be an iterative process which requires the analysis team to deliberate over several generations of situations to achieve an adequate appraisal.

d. Questionnaire

The questionnaire method of analysis is another qualitative technique. As generally described by the literature, a questionnaire is a list of questions developed to elicit responses concerning the security of a system, or any aspect thereof. It is distributed to a group of stakeholders, completed, and returned for analysis by the security team.

The form of the questions asked can be of two varieties, open-ended and closed-ended questions. Fowler provides in-depth discussion on the impact and merits of the type of questions that should be utilized for any given situation (Fowler, 1984, p.87). Generally, open-ended questions are most appropriate for gathering information about system vulnerabilities and threats. They provide a forum for the respondent to express his true opinion or allow the respondent to venture into areas of specific concern to him. This technique has several advantages. The depth and detail of feelings provided by respondents to open-ended questions provides the analyst with a strong sense or perception of the disposition and mood of the environment in which the system must function. When appropriately accumulated and appraised collectively, the questionnaire can provide insight not achievable by other methods. (Patton, 1980, p.28)

In terms of disadvantages, open-ended questions frequently will return topics beyond the desired scope of the analysis, are difficult to interpret, and are nearly infeasible to quantify. Regardless of the type of questions utilized, the questionnaire is also limited by the writing skills of the respondent, the effort required by the respondent, and the inability of the analyst to further address or contemplate responses of particular interest. Although these problems can be significant, they also can be minimized by utilizing established design principles for questionnaires or employing a specialist to design and analyze each questionnaire.

To enhance the decision portion of questionnaires, many practitioners utilize the Delphi technique. The Delphi technique is a survey method that acquires a decision from group consensus. A survey is distributed to a panel of experts involved with a decision. The experts return their responses to an analyst, who accumulates the data, consolidates the results, and redistributes the group results to the respondents. Without communicating among themselves, the respondents reconsider their opinions based on the group opinion and submit their new answers to the analyst. (Hutt, 1988, pp.24-25)

Popularized by the RAND Corporation in the early 1970's, applications of the Delphi method extended to numerous disciplines. While critiqueing the Delphi method, Sackman acknowledged, "It became virtually indistinguishable from questionnaire techniques." (Sackman, 1974, p.5) Nevertheless and in spite of its success, he recommended that it not be used as an alternative to more rigorous, scientific techniques such as questionnaires until it could be thoroughly validated. (Sackman, 1974, p.74)

Although we did not consider the Delphi method as an instrumental component of questionnaire methods, we did make several other assumptions for the comparison evaluation:

- The questionnaire is or was properly engineered by a specialist for our purposes.
- The writing skills of the respondents were not a detrimental factor.
- A sufficient number of questionnaires would be returned to analyze.
- A specialist is employed to assist in the analysis of the questionnaire answers.

Although the assumptions applied to our idealistic questionnaire would certainly have an impact on the values attained by the metrics, they are reasonable and necessary for the purposes of making an effective comparison.

3. Intuitive Predictions

Some results of an evaluation should be predictable. For instance, the checklist method should rank relatively lower than quantitative methods (ALE) in the adaptability criterion because checklists tend to be inflexible and difficult to modify. Predictability provides an intuitive method of determining the usefulness of our technique in lieu of empirical data. Although intuitive predictions cannot provide a fully acceptable validation, they provide a reasonable expectation of correctness from which to justify the collection of data for a more refined validation process.

Using this approach, intuitive predictions were made for each of the criteria. The intuitive predictions provided below are based upon the knowledge gained during the course of our literature review. Each of the seven criteria were individually and comparatively considered with respect to the four methods. For each criterion, the best

and worst method displaying the qualities of the criterion were selected. For instance, each method (quantitative, checklist, scenario, and questionnaire) were examined as a tuple for consistency, then useability, and so forth. Justification for the intuitive predictions that we made are contained in the paragraphs below. Table 4 provides a summary of the predictions.

TABLE 4. Predictions

		PREDICTION
Consistency	High	Checklist
	Low	Scenario
Useability	High	Checklist
	Low	Questionnaire
Adaptability	High	Scenario
	Low	Checklist
Feasibility	High	Checklist
	Low	Questionnaire
Completeness	High	Scenario
	Low	Checklist
Validity	High	Scenario
	Low	Checklist
Credibility	High	Questionnaire
	Low	Checklist

The checklist would be the most consistent of the methods because a checklist's consistent terminology prevent significantly different results as different analysts employed the technique. The questions themselves would remain unchanged and only the differing perspectives and experience of the analysts would cause a difference

in the results. Scenario, on the other hand, was the least consistent of the four methods to be examined. This was justified by the fact that each scenario considered is strongly influenced by the biases of the analysts involved. Hence, with each change in analysts, the scenarios considered will differ significantly.

Because of its unchanging (except for minor modifications) interface with the user, checklist was determined to be the most useable of the four methods. Assuming that it has been appropriately selected to apply to the system it will evaluate, the checklist is easy to pick up and implement. The output is simple to interpret; a vulnerability is highlighted by a lack of a check in the box. Compiling the vulnerabilities into categories is relatively simple and provides an understandable view of system weaknesses. Conversely, the most difficult method to use (the least useable) was questionnaire. While the questionnaire interface does not change from iteration to iteration, the compilation process is very complex. After the questionnaires are returned, the opinions of each respondent must be integrated with all others to derive action recommendations. With the wide variety of opinions that an analyst is likely to receive, this task could be very difficult.

Among the four methods, the most adaptable was the scenario method because its free-form style provides the analyst the options of:

- selecting the approach and analytical procedures that he will employ.
- determining the level of detail that his team will take.
- deciding in what phase of the system life cycle he will conduct his analysis.
- adjusting his approach for each vulnerability or asset he considers.

Alternatively, the least adaptable method was the checklist because the checklist is either designed for a particular system configuration or is designed to suit a very broad class of systems. Universal checklists tend to be fixed in its structure, hence, they are not easily modified to accommodate system changes. In contrast, many of the explicit checklists accommodate change by providing a facility to remove obsolete items and add additional items as the system evolves. Although this is sufficient for modifying the checklist for the system it was intended, the same checklist is not necessarily adequate for another configuration.

While checklists have little ability to adapt, it was ranked the highest among its peers in feasibility. This prediction was based on the fact that the data required to accomplish the checklist was relatively easy to obtain and was the most easily gathered. We considered questionnaire to be the most infeasible of the four methods because of the amount of effort required to prepare, distribute, gather, and compile the necessary data. To prepare an adequate questionnaire, specialists should be involved, making the process longer and more intricate. Once completed, the questionnaires must be compiled into specific action recommendations, a process that will involve deliberations to determine which opinions hold the most merit, and how much weight to place on each. All of these factors implicate questionnaires as the least feasible of the methods considered.

The least likely of the methods to be complete is the checklist method. Our reasoning was reinforced by the fact that checklists are normally designed for a particular system configuration (centralized data processing, distributed processing, etc.) and have only a limited ability to cover a spectrum of systems. Additionally, checklists normally

are designed to check the existence of a particular safeguard, and therefore do not consider all of the elements or relationships that are an integral part of risk management. As such, they lack completeness. Because scenario is a free-form method of analysis, it can consider any aspect of the system that the analyst desires. Hence, system configuration, environment, and all other segments of the system can be considered. For this reason, the scenario method was chosen to be the most complete of the methods.

Validity, the measurement of how closely the process and its results represent reality, was most present in scenario methods of risk management. Scenario was considered to be the most valid because the team of experts involved in the scenario process would be very familiar with the system and its vulnerabilities. Although it is likely that the team would be influenced by their personal biases, scenario was perceived as having the ability to represent the system in terms that the decision-makers would understand and accept as authentic.

Conversely, the output of a checklist was believed to represent the least accurate statement of the actual state of the system. Although this output represents a set of safeguards that are not present, the set of vulnerabilities considered might not have been complete. If a particular vulnerability is not on the checklist, the need for a safeguard might not be addressed, and consequently, it will not be part of the checklist output whether present or not. Alternatively, other methods provide a means of introducing vulnerabilities, the checklist lacks a stimulation mechanism that would prompt concern about a vulnerability not on the list. The analyst must resort to additional methods to determine if his system is vulnerable to a threat that wasn't addressed by the

checklist. Consequently, checklists were not considered to reflect the true state of system security, but rather, the true state of the system that applies to the checklist.

Of the four methods for risk management considered, questionnaire was the most credible. This premise was based on the procedures contained in the questionnaires themselves. Questionnaires collect all of the opinions and impressions of the users of the system itself and as a result, the output (as a product of all of those opinions) should be the most believable to the analyst. On the other hand, checklist methods was the least credible as a result of their inherent inflexibility. Since they are so broad in their extent, they tend to loose their authenticity, particularly in the eyes of the decision-maker.

E. CONCLUSIONS

The evaluation was conducted by the authors. The scores for each methodology in this evaluation are provided in Appendix C and summarized in Table 5. Each methodology was considered individually and entirely prior to proceeding to the next methodology. Because the methods were analyzed without regard to a specific situation, a consensus of our opinions was obtained prior to responding to each metric to ensure we applied the metrics consistently. The following paragraphs are the conclusions drawn from the results of the evaluation process.

1. Observations

a. Strengths and Weaknesses of Methodologies

Close scrutiny of Table 6 and Appendix C reveals that the strengths and weaknesses of each method are easily discernable. For instance, ALE ranked the highest

TABLE 5. *Comparison of Predictions to Results*

		PREDICTION	MEASUREMENT
Consistency	High	Checklist	Checklist
	Low	Scenario	Scenario
Useability	High	Checklist	Checklist
	Low	Questionnaire	Questionnaire
Adaptability	High	Scenario	Scenario
	Low	Checklist	Checklist
Feasibility	High	Checklist	Scenario
	Low	Questionnaire	ALE
Completeness	High	Scenario	Scenario/Questionnaire
	Low	Checklist	Checklist
Validity	High	Scenario	Scenario/Questionnaire
	Low	Checklist	Checklist
Credibility	High	Questionnaire	Checklist
	Low	Checklist	Questionnaire

in credibility and lowest in feasibility within its own criteria. Both results can be attributed to the objective nature of the method. ALE is dominated by numerical values which lend credibility to the procedure (Quade, 1975, p.161). Because numbers are manipulated so extensively, and the detail required to accomplish the risk analysis requires an enormous amount of data, the feasibility of a complete analysis is low. Instinctively, checklists are extremely useable while not very complete because of version control. Scenarios are strongest in their ability to adapt to the situation or circumstances, an attribute that makes consistency difficult to maintain. And finally, questionnaires are most complete because of their ability to survey large numbers of people with an

unlimited range of responses and topics, but holds little credibility due to the lack of intuitiveness of its technique.

This ability to recognize the strengths and weaknesses of a particular methodology can be a useful tool to analysts who are developing hybrid procedures. Ideally, an analyst could select a method that demonstrates a significant strength under a specific criterion and combine it with another method that demonstrates an entirely different strength to achieve a hybrid methodology that optimizes the strengths of each method in a particular situation. The implications of designing a custom methodology are imaginable.

b. Holistic Comparison Capability

The evaluation method presents a holistic comparison of dissimilar methodologies - a capability not claimed by any other technique to date. The scores achieved from the metrics by a method collectively reflect the presence of attributes in that methodology. The comparison therefore is made between the scores for attributes, or criteria, and represents the relative strengths or weaknesses of that method. This means the comparison of *Consistency*, for example, is considering the consistency of each methodology individually and provides a holistic view of all risk management methodologies under evaluation.

Using a holistic view of risk management methodologies, dissimilarities are no longer an impediment to comparison. Two examples that have presented difficulties for comparison in the past were the use of quantitative data versus qualitative data and risk analysis versus risk management (the former lacking a decision-making

process, while the latter includes that process). Both of these problems are overcome by this holistic comparison capability.

TABLE 6. Method Evaluation Results

	ALE	CHECKLIST	SCENARIO	QUESTIONNAIRE
CONSISTENCY	**.625**	**.750**	**.250**	**.625**
Reliability	.500	1.000	.500	.500
Consistent Terms	.750	.500	.000	.750
USEABILITY	**.579**	**.817**	**.567**	**.546**
Error Handling	.666	.666	.666	.333
Simple	.500	1.000	.250	.750
Ease of Use	.750	1.000	.750	.500
Understandable	.400	.600	.600	.600
ADAPTABILITY	**.625**	**.250**	**.750**	**.500**
Portable	.750	.000	1.000	1.000
Modifiable	.500	.500	.500	.000
FEASIBILITY	**.333**	**.546**	**.694**	**.472**
Availability	.333	.666	1.000	.333
Practicality	.000	.750	.750	.750
Scope	.666	.222	.333	.333
COMPLETENESS	**.500**	**.185**	**.722**	**.722**
Scope	.666	.222	.333	.333
Elements	.500	.333	1.000	1.000
Element Attributes	.333	.000	.833	.833
VALIDITY	**.444**	**.435**	**.583**	**.583**
Relevancy	.666	.333	.666	.666
Scope	.666	.222	.333	.333
Practicality	.000	.750	.750	.750
CREDIBILITY	**.639**	**.666**	**.472**	**.367**
Intuitiveness	.777	.333	.444	.222
Reliability	.500	1.000	.500	.500

2. Confirmation of Predictions

The real significance of this evaluation model is in the comparison of the criteria and attributes across each methodology simultaneously. Figure 8 graphically illustrates this comparison between methods. The most distinctive methodology for any particular criterion can be selected by comparing the scores horizontally. Although this may seem like a minor achievement, when the reader considers that each of the methodologies have been evaluated by precisely the same gauge - a standard - the results achieve a level of credibility not previously attainable. The credibility of our evaluation model (not the criteria of credibility) is further substantiated by the general correlation of our evaluation with the predictions (see Table 5).

As expected, of the four methods evaluated for useability, checklists ranked the highest, whereas, questionnaires ranked the lowest. The underlying reasons for these differences can be ascertained by a closer examination of the metrics. Checklist, with its simple and easy to use interface, its ability to quickly and efficiently recover errors, and its inherent understandability caused it to rank highly in useability. Questionnaire, alternatively, is relatively simple to administer, but is not easy to use because of the complexity introduced when questionnaires must be compiled. Questionnaire also ranked poorly because of the difficulty in finding and correcting errors in the morass of data received in the responses.

Another example of the metrics corresponding with the prediction is the criterion completeness. The most complete methodology (as determined by the evaluation method) is scenario or questionnaire while checklist rates last of the four. Although the

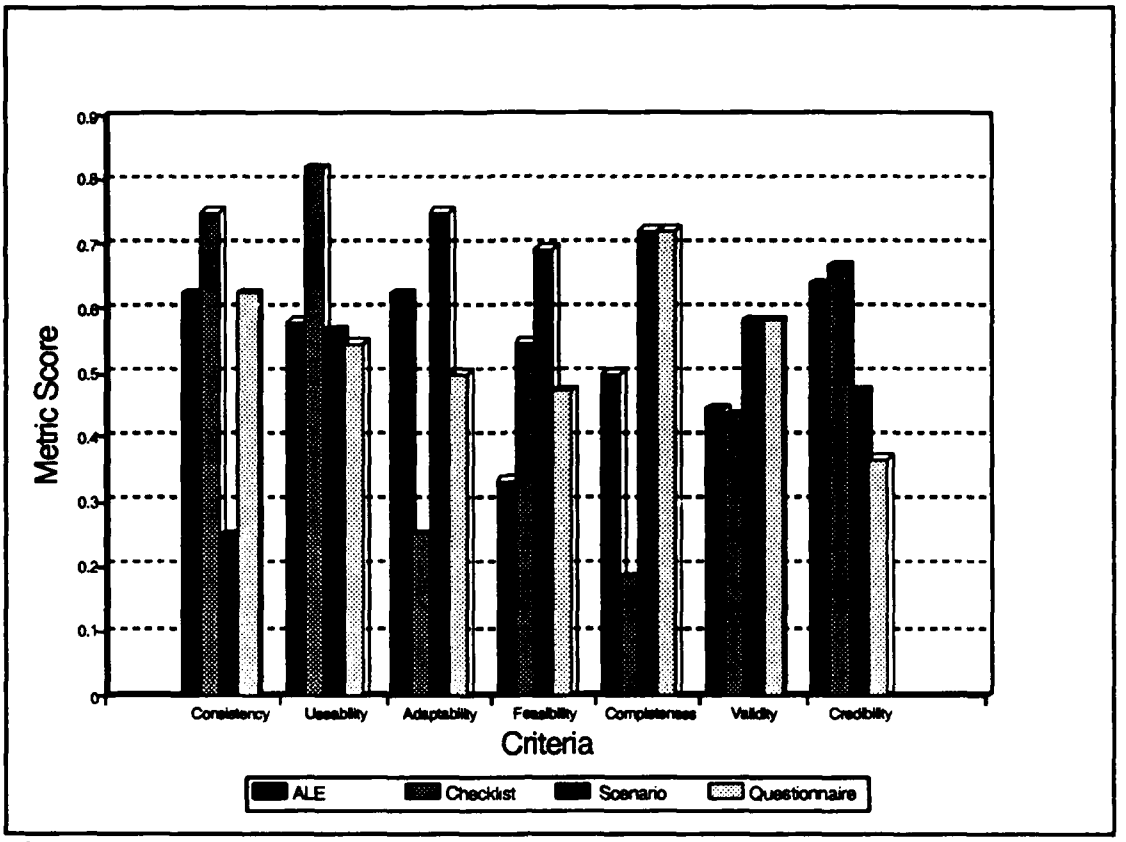

Figure 8. Methodology Criteria Values

data collection is different for each of the methods, scenarios and questionnaires collect input data which is limited only by the experience and knowledge of the participants. Completeness is achieved by the open-ended approach to data collection (free-form scenarios or open-ended opinion type questions). On the other hand, checklists are closed-ended in their data collection. Input is limited to only those items contained on the checklist. As a result, checklists tend to be incomplete.

Similar arguments could be made to explain the close correlation between the predictions and measurements of the criteria of consistency, adaptability and validity.

Admittedly, the converse of the above discussion is also present. Several examples could be referenced from the results listed in Tables 5 and 6 that don't correspond to our predictions. In fact, the criterion credibility provides a complete contradiction to our predictions.

The purpose of analyzing our results in the context of our predictions is to provide an approximation of the usefulness and integrity of the metrics. The strength of the approximation is based on a preponderance of the results from the application of our metrics correlating with the intuitive predictions. This is an indication that the metrics are measuring, to some degree, the characteristics that we have determined to be instrumental in a comparison process. While an incomplete correlation provides sufficient evidence that a problem does exist, it also implies that the metrics are providing a viable technique for discerning methodology characteristics. A perfect correlation or a lack of any correlation would have indicated that the metrics were worthless or ineffectual, respectively.

The few results that differ significantly from the predictions in our evaluation can be interpreted as one of four possibilities:

- The metrics are not contributing to the measurement of a particular attribute.
- There are an insufficient number of metrics to fully measure the presence of an attribute.
- The prediction is too imprecise to provide a distinguishable subject for evaluation by the metrics technique for that criterion.
- A combination of the problems above exists.

Although the comparison of our evaluation results against our predictions is adequate for approximating the usefulness and integrity of the metrics, it is scarcely an adequate technique for determining the nature and extent of the above problems. Any conclusion drawn from this evaluation about the problems which exist would be purely speculative. The task of validating the metric evaluation technique is deferred to more empirical and analytical methods and will be discussed in Chapter V.

In essence, this chapter confirms the metrics evaluation technique is providing an acceptable, standardized measurement of a methodology's attributes upon which to base a more sophisticated comparison of risk management tools. Additionally, it is expected that a refinement of and additions to the metrics developed in Chapter II will be necessary. Further discussion of this topic will be presented in Chapter V.

3. Applicability to Hybrid Methods

A feature of the metric evaluation technique that should be emphasized concerns the comparison of hybrid methodologies which are representative of most risk management tools. The term hybrid is used to denote methodologies that employ a combination of different risk management methods. Theoretically, they incorporate the strength of each method in a complementary fashion, therefore, they would obviously rate higher on most criteria against the generic methods we utilized. The premise to use generic and discrete methodologies for this application of the evaluation metrics was to insure our results were not erroneously derived from the impact of a hybrid method.

It was imperative that the application of the metrics reflected a logical basis and relatively accurate results, as described in the previous section, before considering a

more complex application against hybrid methods. This chapter has achieved that requirement. Consequently, Chapter IV submits the evaluation model to one institutional and three commercially available risk management tools which are hybrids.

IV. A PRAGMATIC COMPARISON OF RISK MANAGEMENT TOOLS

A. INTRODUCTION

In Chapter III, the metrics were applied to the evaluation and comparison of generic risk management methodologies. That procedure provided a rudimentary indication that the metrics indeed measured certain attributes, and consequently criteria, which are considered essential to the evaluation and comparison process. However, the methodologies examined were distinct and independent of each other. Even though the capability to evaluate and compare individual methods is beneficial, the significance of the metric evaluation is in its application to hybrid methodologies. Hybrid methodologies are representative of the majority of tools that are currently available to computer security risk managers. This chapter demonstrates the strength of the metrics evaluation technique by evaluating and comparing a small sample of four hybrid tools. The analysis was conducted by the authors.

1. Prerequisite Computer System Case

Unlike the comparisons made in Chapter III, the comparison of hybrid tools necessitates using a tool within the context of a test case computer system. The evaluation in Chapter III measured attributes of the methodologies (not the mechanics involved) to determine the presence of correlations between observations and predictions. This chapter applies the metric evaluation technique to hybrid tools for the purpose of

determining the suitability of their use for a specific computer system configuration, by a particular analyst/manager.

The mechanical process required by each tool is a significant part of the evaluation. It must be performed to the extent that sufficient information has been obtained by the evaluator to make a valid response for several of the metrics. To ensure a consistent comparison, each tool must be provided with an identical scenario. For the purposes of this thesis, this is accomplished by a computer system case which can provide a consistent source of data for:

- Assets, values, and procedures.
- Relationships, configurations, and environmental circumstances.

A narrative of the computer system case used in this chapter is provided in Appendix D.

2. Disclaimer for Case System

The case system presented in Appendix D is modelled after an actual computer system laboratory provided by the Naval Postgraduate School's (NPS) Administrative Sciences Department. The information provided in this case system should not be construed to represent actual conditions, procedures, or circumstances of any kind that may exist at the Naval Postgraduate School. The laboratory was used only to provide a framework for the development of the fictitious system in Appendix D. Additionally, no evaluation nor statement is made or implied concerning this laboratory. The case system is provided to ensure consistent inputs for the evaluation of the tools.

B. RISK MANAGEMENT TOOL DESCRIPTIONS

In Chapter I, it was established that a plethora of tools exist. In fact, more than 30 automated risk analysis and risk management packages are commercially available (Gilbert, 1989). However, the cost of these packages has prevented the purchase of any risk management tool for this study. The prices range from $350 to $29000 (Johnson, 1987, p.36). The lack of travel funds and limited time available further prevented the use of the NIST Risk Management Research Laboratory for conducting an evaluation and comparison. Consequently, demonstration versions of three automated risk management tools, with their accompanying documentation (which consisted mainly of advertisement literature) were acquired. Although extremely limited in their capability to perform all the functions provided by the actual tools, they proved to be sufficient to provide the necessary examples upon which to base answers to the metrics. A fourth risk management tool, a manual survey, was included in the evaluation to demonstrate that the metrics' comparison capability is not limited exclusively to automated tools and that dissimilar tools can be compared on an equal basis.

Each of the following sub-sections provides general background information and a brief description of the packages utilized in this evaluation. Narrative depictions of these tools have been written devoid of any preconceptions introduced by advertisements concerning the features and underlining methodology of the products. This should provide the reader with a factual perception of the tools from which to consider the difficulty of a comparison for suitability without metrics.

The order of presentation was arbitrarily selected and does not reflect any preference or endorsement by the authors.

1. RiskPAC

RiskPAC is a software product that facilitates the application of qualitative questionnaires and production of reports. The questionnaires are predeveloped for categories of analysis concerns, such as physical security, personal computers, telecommunications, and computer system applications, each of which are purchased separately.

Regardless of the questionnaire selected, answers are selected from those responses provided. The results of a questionnaire are stored in survey format for later access by the user for modification or correction. RiskPAC System Manager provides a customizing capability for questionnaires that allows a user to tailor questions to his specific situation.

Several reports are provided that reflect a general risk category, a risk level, and a risk description. The risk level is represented by a number between one and five inclusively, with five being the worst case. The method of computation for the risk level is undisclosed to the user. RiskPAC is an interactive, microcomputer based risk assessment package. Decision analysis is possible by weighing and scoring selective questions.

Requirements include an IBM or compatible XT or AT and 256KB of memory. Two diskette drives are necessary if a fixed disk drive is not available. The

interface is menu driven. A user's manual is provided, and training is provided upon request.

The vendor of RiskPAC is Profile Analysis Corporation, Ridgefield, Connecticut, a subsidiary of Computer Security Limited.

2. Naval Postgraduate School (NPS) Security Survey

This is a manual questionnaire provided to the Automated Data Processing (ADP) security officer or the project manager for all Automatic Information Systems (AIS) at NPS. The survey collects and documents:

- Basic Data - Information concerning system identification, system location, responsible individuals, functional purpose, data characteristics and value.
- System Description - Hardware components, operating mode, value, and software inventories.
- Risk and Vulnerability Assessment - Countermeasure and Vulnerability relationships.

The risk and vulnerability assessment portion of this survey implements a checklist in the form of a questionnaire which allows the user to depart from the traditional "yes" or "no" style of checklists to a broader range of responses.

At the completion of a category of vulnerability versus countermeasure, the user makes a subjective evaluation of the risk for that category. Footnotes are used to prevent ambiguous terminology, and an attached supplement provides narrative examples of vulnerability descriptions that may be used in documenting deficiencies.

3. RiskCalc

RiskCalc is an automated package with spreadsheet-like characteristics. It has no underlying methodology, but a shell provides the purchaser with the means to create any risk assessment computational model that is desired. The user answers a questionnaire which inputs asset value and probability information to the system. Optional changes to the system variables provides an immediate "what if" capability for determining the most cost effective safeguard.

Demonstration models are available which can be used as pre-designed questionnaires or custom tailored to fit the users requirements through the use of the System Manager.

Requirements include an IBM or compatible PC, XT, or AT and 512KB of memory. A fixed disk drive is recommended but not required. The interface is menu driven with an on-line help facility. A user's manual and a system administrator's manual are provided. A one day, on-site training session is provided with the purchase, and a three-day risk management course will be provided upon request.

RiskCalc was developed and is distributed by Hoffman Business Associates, Inc., Chevy Chase, Maryland.

4. Los Alamos Vulnerability Assessment (LAVA)

LAVA is a quantitative and qualitative, automated vulnerability assessment package with an underlining methodology that combines questionnaire, scenario and checklist techniques. It provides identification of missing safeguards through an extensive

group of questionnaires. The threat environment is characterized by three broad categories:

- Natural Hazards.
- On-Site (Direct) Humans.
- Off-Site (Indirect) Humans.

LAVA stresses a group consensus approach to answer the questionnaires. The group ideally is composed of personnel from a broad spectrum of backgrounds and expertise of the system under analysis.

Requirements include an IBM or compatible XT or AT and 512KB of memory. A fixed disk drive is recommended, although, not required. The interface is free-form. A user's manual is provided; however, training at a hands-on workshop which lasts a full week is required.

LAVA was developed and is distributed by Los Alamos National Laboratory, Los Alamos, New Mexico.

C. EVALUATION AND COMPARISON

The evaluation was conducted by the authors for the purpose of demonstrating the use of the metrics on hybrid tools and not to rank one tool better than another. We are obliged to remind the reader that the results of this evaluation are based on demonstration versions of the three commercial products, and not the complete products. This evaluation is not intended to provide an unfair advantage to any commercial vendor nor should it be construed as an endorsement or condemnation of the subject tools.

1. Procedures

The evaluation process followed four simple steps:

- Execute the demonstration program for one of the tools (in the case of the manual survey, it was conducted as prescribed by the instructions).
- Answer all questions as accurately as possible, providing data in the context of the case computer system.
- Upon completion of the data collection, review all reports that were available from the demo version and accompanying literature.
- Immediately following the case study demonstration, the metric questions are answered, and values are calculated.

As the first tool completed its entire evaluation, including the compilation of the metrics' values, the next tool was processed through the above procedures. This evolution continued until all four tools were evaluated. While answering the metrics, occasionally quick reviews of pertinent portions of a demonstration were conducted to resolve indecision concerning an answer.

The Normative Model as described in Chapter III was utilized consistently throughout this evaluation process.

2. Performance of the Tools

It became readily apparent that the demonstration software and its accompanying literature were sufficient to conduct this evaluation. Although RiskPac provided a "story book" demonstration which prevented the use of data from the case computer system, the presentation provided sufficient and explicit examples upon which to base metric decisions. We encountered no difficulties in answering the metric questions for any of the tools evaluated.

The tabulated results of this evaluation are found in Appendix E and are summarized graphically below in Figure 9. Since the focus of the effort in this chapter is to demonstrate the application of the evaluation procedure, no attempt is made to explain or justify a particular grade earned by a specific tool; doing so would divert the reader's attention from the evaluation procedure.

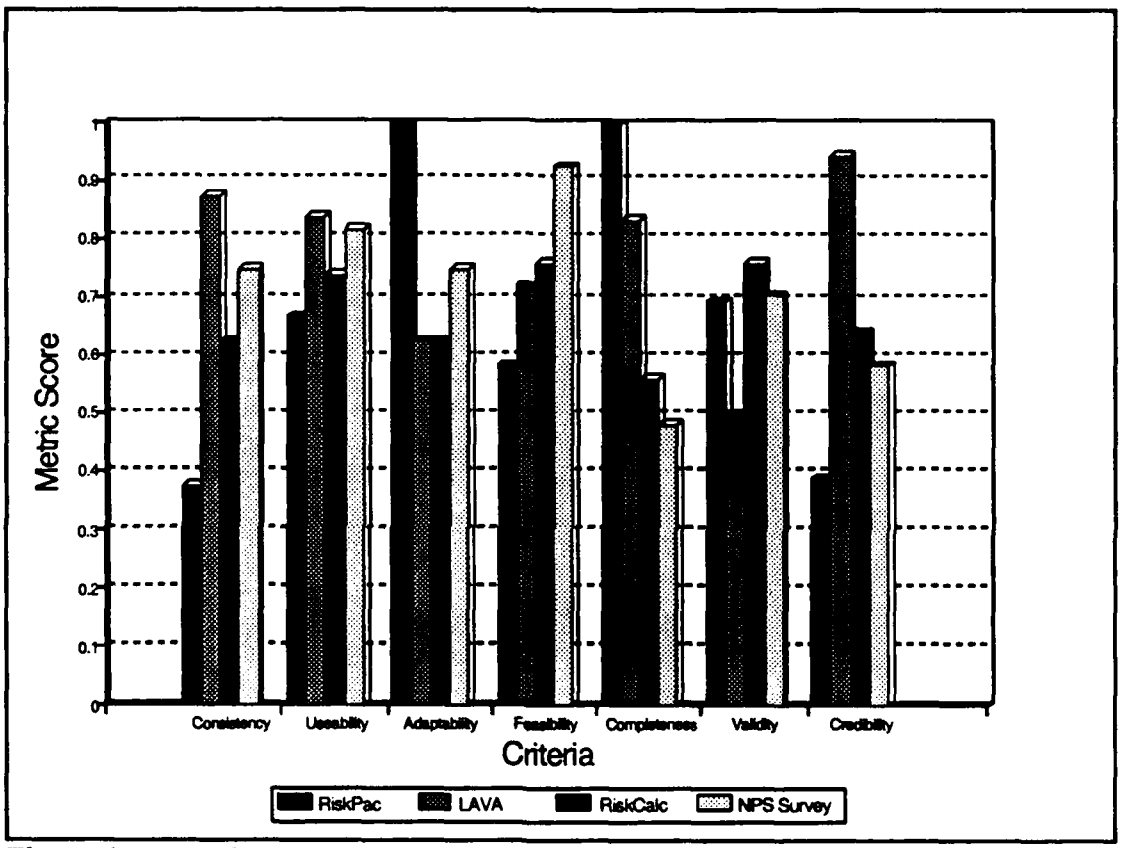

Figure 9. Tool Criteria Values

3. Performance of the Metrics

Discussion of the evaluation results can be focused on the three perspectives introduced in Chapter III. Those perspectives, reiterated in the context of the evaluation conducted in this chapter, consist of examining the results of each tool separately,

examining the results of each tool in comparison to each other, and finally, examining the results by comparing the suitability index of each tool.

The metric evaluation technique, demonstrated in Chapter III, indeed measures the criteria that describe the suitability of a risk management tool. Selecting any one of the tools (RiskPac, LAVA, RiskCalc, or the NPS ADP Security Survey) from Figure 9, the reader can numerically rank each of the criteria for that specific tool and determine its strength or weakness.

One of the primary advantages of the metric evaluation technique is its ability to compare a single criterion across several methods or tools at one time. This was aptly demonstrated with dissimilar methodologies in Chapter III. However, the evaluation conducted in this chapter demonstrates the metrics' ability to discern criteria between similar as well as dissimilar tools. As an example, examining the criterion 'completeness' in Appendix E, shows that RiskPac is most complete, followed by LAVA, next RiskCalc, and finally, the NPS ADP Security Survey. These results indicate that the manual survey, when compared to three automated packages, presented no significant problem for the evaluation, nor was there any difficulty in discerning diffcrences between the criteria of the three automated tools.

Utilizing the Normative Model to develop a 'suitability index', a broader perspective that includes all the criteria is possible. This technique gives equal consideration to all the criteria (criteria weights are equal to one) and therefore, provides a holistic ranking all of the tools. The 'suitability index' for the sample risk management tools used in the evaluation for this chapter are presented in Figure 10.

Although not a critical problem for the selection of tools from the samples evaluated in this chapter, the 'suitability index' on some occasions may fail to numerically discriminate between tools. For instance, had the evaluation consisted only of RiskPac and RiskCalc, then a decision based on the 'suitability index' would not be possible because of equal values. As an alternative, the evaluator could resort to selecting a tool based on a criteria comparison.

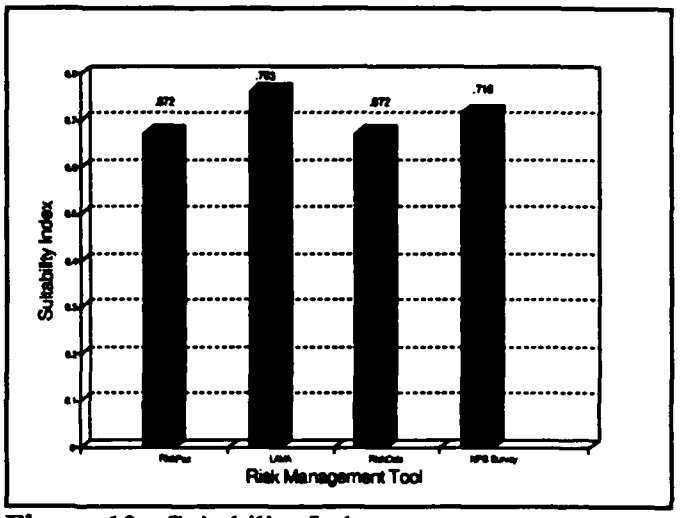

Figure 10. Suitability Index

In conclusion, this chapter has demonstrated the practical use of the metrics developed in Chapter II for evaluation and comparison of hybrid risk management tools. Utilizing the Normative Model presented in Chapter III, the selection of the most suitable tool has been simplified for the case computer system described in Appendix D. Furthermore, had the need existed to make a selection based on a more discriminating analysis, that capability was available by examining and comparing criteria values or attribute values of particular concern.

V. OBSERVATIONS, RECOMMENDATIONS, AND CONCLUSIONS

A. OBSERVATIONS

1. Metric Formulation

Making the right choice of which risk management tool from among a wide variety of commercially available packages is difficult. The complexity of the decision is further perplexed by the multitude of approaches used as the basis for those packages. In this study, we utilized a systematic approach to compose a description of the suitability of a particular risk management method by defining suitability in terms of its underlying criteria. The attributes that compose a criterion were then defined and expressed in terms of metrics. By progressively decomposing suitability into metrics, comparisons are readily made. The comparisons facilitate making a choice of risk management tools. The multiple criteria evaluation method was applied against four underlying methods for conducting risk analysis to test the procedure, and then was employed to evaluate four readily available risk management packages.

2. Multiple Criteria Evaluation Method Application

a. Risk Management Methodologies

(1) Feasibility Criterion. In making intuitive predictions about the metric evaluation of the underlying methods in Chapter III, we came to our conclusion that questionnaire would rank lowest amongst the methods because of the inherent

difficulty in collecting and compiling the data provided in the questionnaires. This underlying difficulty should have a significant negative impact on feasibility of the questionnaire method. At the same time, the simplicity and ease of use found in the data collection of the checklist method should make the method more feasible despite a certain amount of difficulty associated with compiling the results into a decision. For these reasons, our intuitive predictions for feasibility were that checklist would rank highest and questionnaire would rank lowest. When Template 1 of Appendix C is examined, however, it is seen that checklist ranked second (.546 compared to scenario's .694) and questionnaire ranked third (.472 compared to quantitative's .333).

To determine the source of the discrepancy between the predictions and the resultant values, it was evident that the metrics, as written, capture the feasibility of only the data collection process. Our predictions, on the other hand, considered all aspects of the process, including the collection, processing, and interpretation of the data, pointing out that an aspect of the process was overlooked when the metrics were generated. The question is: "Are the metrics versatile enough to incorporate the compilation and interpretation processes with data collection?"

As test of the ability of the metrics to accommodate a modification of the definition of a criterion, we will re-write the definition of feasibility as: "The required data is available, and can be economically gathered, processed, and interpreted." To measure the feasibility of all aspects of the process, the attributes of *availability, practicality,* and *scope* remain valid as measures of data collection. The attributes of *ease of use* and *simplicity* (as explained in Chapter II) are added to provide a measure of

processing and interpretation and a new attribute called *processability* is added to measure how readily data is processed by the method. Processability involves two metrics to measure the ability to economically compile and interpret the collected data. The two metrics are described as:

- "Does the process facilitate the compilation of data in a form that is useable by the analyst and decision-maker?"
- "Is the compiled data in a form that is readily interpreted by the decision maker?"

When the one new and two previous criteria are added to the measurement process, the results are more in keeping with the predictions made earlier; checklist ranks highest, quantitative second, scenario third, and questionnaire lowest (see Table 7).

TABLE 7. Feasibility Criterion for Risk Management Methods

	QUANTITATIVE	CHECKLIST	SCENARIO	QUESTIONNAIRE
FEASIBILITY	.542	.690	.513	.444
Availability	.333	.666	1.000	.333
Practicality	.000	.750	.750	.750
Scope	.666	.222	.333	.333
Simple	.500	1.000	.250	.750
Ease of Use	.750	1.000	.750	.500
Processability	1.000	.500	.000	.000

If the same new metrics are applied to the commercial tools examined in Chapter IV, the NPS ADP Security Survey would move from its first ranking in feasibility to third. All other tools would remain in their relative positions to each other. Table 8 provides the resulting values for feasibility when the metrics are applied as described above.

TABLE 8. Feasibility Criterion for Risk Management Tools

	RiskPac	LAVA	RiskCalc	NPS
FEASIBILITY	**.708**	**.819**	**.838**	**.796**
Availability	.000	1.000	1.000	1.000
Practicality	.750	.500	.500	1.000
Scope	1.000	.666	.778	.778
Simple	1.000	.750	1.000	1.000
Ease of Use	1.000	1.000	.750	1.000
Processability	.500	1.000	1.000	.000

While we are not advocating wholesale modification of the criteria, attributes or metrics, it is valuable to know that if a criterion is found to be measuring something other that what was intended, a modification can be accomplished readily.

(2) Credibility Criterion. Probably the most interesting results of the use of the metrics against the underlying methods was that of credibility. As a result of the evaluation of the four risk management methods, checklist was found to have the greatest value (.666) while questionnaire had the lowest (.367). A cursory examination of Table 5 indicates that a complete reversal of the predictions took place in the measurements.

Initially, we were prepared to accept the possibility that the metrics were inappropriate (i.e. they were measuring the wrong thing), or else the metrics were incomplete in capturing the criterion of credibility. However, upon detailed analysis of the predictions and the measurements, we determined that a discontinuity existed in the levels at which the two processes (prediction and metric measurement) were conducted.

The predictions were based on an in-depth understanding of each of the methods ultimately evaluated, albeit greatly generalized. In spite of this understanding, our literature review and focus of understanding was oriented to the analyst rather than the decision-maker. When making the predictions, we examined the methods for credibility from the perspective of the user of the system.

The metrics, alternatively, were written from the perspective of the decision-maker rather than the analyst. Intuitively, when the methods are examined from the perspective of the decision-maker, our predictions are significantly different. Questionnaire, because it must compile opinions from personnel with a variety of backgrounds and knowledge, will not hold as high a degree of credibility for the decision-maker when compared to numeric methods (quantitative), processes where the decision-maker's opinion is included (scenario), or processes where an established routine exists (checklist). When examined from the same perspective as the metrics, our intuitive predictions would indicate that checklist would rank best and questionnaire would rank lowest. These intuitive results are in consonance with the results obtained from the metrics as presented in Template 1 of Appendix C.

The experience with credibility in our evaluation of risk analysis' underlying methods illustrates the strength of the metric evaluation method in forcing a consistent level of examination across all levels of the evaluation.

b. Risk Management Hybrid Tools

One of our greatest concerns as the metrics were developed was the need for the metrics to provide enough detail in the evaluation to facilitate discrimination

between commercial risk management packages that would share many of the same characteristics. After an examination of the results of the evaluation conducted in Chapter IV, we determined that there was sufficient differentiation between packages to make a suitability determination.

In some cases, the numeric separation between packages for attributes and criteria was negligible (as a result of the small number of metrics that comprise an attribute). For example, is RiskPac really significantly more modifiable than the NPS survey because it scored 1.000 to the NPS survey's .500? Possibly not, because the difference between the two scores is a single "yes" to one metric.

However, the significance of the suitability decision is not influenced by a small number of metrics. A clear decision can be made at the criteria or suitability level, as long as some preselection threshold is established by the organization doing the evaluation.

3. Multiple Criteria Evaluation Method Role

The procedure has proven to be flexible and useful. The multiple criteria method of evaluation has demonstrated its flexibility by functioning in a predictable manner against both risk management methods and risk management tools. Most significantly, the multiple criteria evaluation technique described in previous chapters provides the analyst and decision-maker with a useful tool for culling relevant information about risk management packages. Patton describes the importance of utility in evaluations:

> In studying the utilization of evaluation research I found that decision makers and information users did not expect evaluation reports to produce "truth." Nor did they treat evaluation reports as containing "truth" in any fundamental sense. Rather, they viewed evaluation findings as additional information that they could combine with other information.... The purpose of evaluation research, then, is to provide relevant and useful information to decision makers.... (Patton, 1980, p.273)

The usefulness of a technique that can differentiate between similar commercial risk management applications is momentous. An individual using this technique to make a comparison can rest assured that the metrics provide a consistent and logical application of the criteria to all subjects of the evaluation. Hence, a standard is established and evaluations are dependably controlled.

B. RECOMMENDATIONS FOR FUTURE RESEARCH

Given the potential for applications of the multiple criteria evaluation technique, the following research areas appear to be promising avenues for the future.

1. Multiple Criteria Evaluation Method Refinement

We have presented a comparison of predictions and observations in Chapter III which provided for continued investigation. However, those results provided no empirical data upon which to determine the validity of the metrics or the technique, nor were they designed to validate the metrics. Further analytical studies should be conducted to make that determination. A requirement to conduct detailed empirical and analytical studies of the performance of the metrics under varying conditions exists to determine their validity and modify or enhance accordingly. Wan addresses this problem:

> Finally, the mathematical results obtained from our analysis will have to be interpreted and compared with observations and empirical data available for the modeled phenomenon. If the results from the model are not sufficiently realistic,

> we must return to the formulation phase of the process and revise the mathematical model. Even if the results are consistent or comparable with observations, we may still want to revise the model and make it more sophisticated by incorporating more questions or more influencing factors. This interpretative phase of the process is sometimes called 'model evaluation' or 'model validation.' (Wan, 1989. p.2)

Another aspect of ensuring the validity of the results reported in this thesis is to perform the evaluation technique using a different evaluation team in an attempt to replicate our results against these and other packages. While different circumstances would produce different scores and rankings than we derived (as is expected with differing evaluation teams), the same amount of differentiation between packages should exist in the results of other researchers. Further research is also necessary to develop additional criteria, attributes, and metrics to improve the differentiation between evaluated packages.

The step-by-step process of performing the evaluation technique lends itself to automation. After the technique has been validated through further experimentation, investigation into automation should be considered. With simple macros on a spreadsheet program, the multiple criteria evaluation method could provide a microcomputer implementation that would provide the following benefits:

- The answers to repetitive metrics would have to be provided only once.
- A choice of mathematical models (normative relationship, significance coefficients, or metric utility) that automate computations could be made easily available.
- Statistical inferences from the scores could be prepared and reported automatically.
- Textual and graphic reports could be generated automatically.

2. Extrapolation of Metrics Approach

The risk management process has three necessary components as depicted in Figure 11. The multiple criteria evaluation method presented in this thesis provides a technique for ensuring that the risk management method is suitable. The risk management team of the organization making the evaluation and the system to be evaluated also play an important role in this environment.

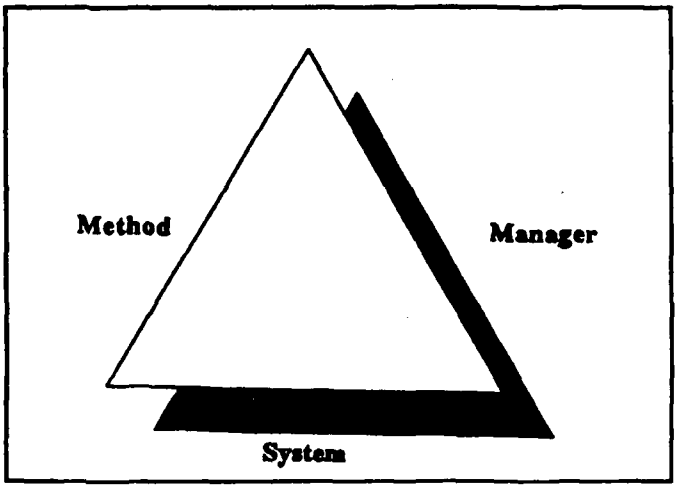

Figure 11. Risk Management Environment

Figure 11 illustrates the triad of elements that comprise the components of risk management. This thesis has introduced and justified the use of metrics to determine the suitability of the methods involved with the risk management process. The other two components, the risk management team and the computer system itself, could also be evaluated utilizing the same multiple criteria techniques introduced in this thesis.

Risk management teams, consisting of analysts and managers, all have varying degrees of competence. Metrics should be developed to measure the competence of the participating team to identify its weaknesses to ensure that the risk management method selected compliments those weaknesses.

At the same time, every system that the risk management team and methods will be employed against will differ in complexity. The computer system could be

quickly evaluated with metrics designed to measure the level of complexity involved in the system.

A risk management method or tool is evaluated for its suitability for and organization's goals and objectives. In the same way, the management team should be evaluated for its competence and the system should be evaluated for its complexity. Together, the three components produce a balance which provides a complete picture of the risk management process for an organization.

3. Optimal Risk Management Method Development

Chapter I discussed the implications of our research regarding the development of an optimal risk management method. After a thorough review of the metrics, attributes, and criteria, we feel that a "best" risk management technique cannot be developed that will satisfy the needs of all organizations. Through the consideration of all aspects of risk management (the method, the management team, and the system) with an evaluation technique similar to the multiple criteria evaluation, satisfaction of a particular organizational need can be fulfilled. Additional research that examines risk management from a holistic viewpoint may be able to classify existing risk management tools to make general categorical statements of their capabilities. For instance, a particular risk management tool might fulfill the needs of a particular group of organizations. An organization's search and comparisons for a suitable tool would be greatly simplified if the comparison were limited to a small sample that was selected from a category of tools determined to be most likely to fulfill their requirements.

C. CONCLUSIONS

This thesis has established a standardized set of metrics in a structured relationship which can be used to evaluate risk management methodologies and tools for their suitability in a given organizational situation. The metrics were successfully applied to four computer security risk management methodologies to develop an informal validation. Finally, the metrics were used to evaluate four hybrid computer security risk management tools as a test of the multiple criteria evaluation method and to demonstrate its use. Its versatility was exemplified by the use of dissimiliar tools. Several suggestions for extension of the concepts developed in this thesis have been provided to guide future research.

There is an ongoing effort by the participants in the Risk Management Model Builder's Workshops to identify the elements and relationships associated with risk management as well as a continued effort to develop automated and manual tools to implement their ideas. As a result, there is a growing need for a consistent and reliable method to evaluate the tools to find the one that best suits the requirements of the enterprise. The multiple criteria evaluation technique described in this thesis can fulfill that need and provide a deeper understanding of the risk management process. The challenge for the future is to exploit that understanding to improve the process in all organizations concerned with computer security risk management.

APPENDIX A. EVALUATION METRICS

TEMPLATE 1

CRITERION: CONSISTENCY (C_1)

ATTRIBUTES	METRIC	YES	NO	VALUE
Reliability (A_1)	1. Does the process provide a mechanism to reduce the introduction of personal bias? (m_1)			
	2. Does the method provide a mechanism that reduces the impact of uncertainty? (m_2)			
	RELIABILITY Attribute Value (A_1)=$\{m_1+m_2\}/2$			
Consistent Terminology (A_2)	1. Is a standard language established? (m_1)			
	2. Are the method's elements defined for the user? (m_2)			
	3. Does the method request input in designated units? (m_3)			
	4. Is the input requested unambiguous? (m_4)			
	CONSISTENT TERMINOLOGY Attribute Value (A_2)=$\{m_1+m_2+m_3+m_4\}/4$			
Metric Value	CONSISTENCY Metric Value (C_1) = $\{A_1+A_2\}/2$			

TEMPLATE 2

CRITERION: USEABILITY (C_2)

ATTRIBUTES	METRICS	YES	NO	VALUE
Error Handling (A_1)	1. Can data entry errors be readily identified? (m_1)			
	2. Does the process facilitate the handling of data entry errors? (m_2)			
	3. Is the process insensitive to insignificant data accuracy errors? (m_3)			
	ERROR HANDLING Attribute Value (A_1)=$\{m_1+m_2+m_3\}/3$			
Simple (A_2)	1. An expert knowledge base is not required to operate the process. (m_1)			
	2. Are complex relationships mitigated for the user? (m_2)			
	3. Is the problem domain well defined? (i.e. is the problem bounded?) (m_3)			
	4. Special training is not required. (m_4)			
	SIMPLE Attribute Value (A_2)=$\{m_1+m_2+m_3+m_4\}/4$			
Ease of Use (A_3)	1. Is there a standardized interface? (m_1)			
	2. Is one iteration clearly differentiated from another? (m_2)			
	3. Is the process well structured and logically sequential? (m_3)			
	4. Is the information requested of the user relevant? (m_4)			
	EASE OF USE Attribute Value (A_3)=$\{m_1+m_2+m_3+m_4\}/4$			
Understandable (A_4)	1. Is the underlying premise explained? (m_1)			
	2. Is the premise comprehendible? (m_2)			
	3. Are the terms unambiguously defined? (m_3)			
	4. Are relationships between elements explained between phases or iterations? (m_4)			
	5. Are decision points clearly identified? (m_5)			
	UNDERSTANDABLE Attribute Value (A_4)=$\{m_1+m_2+m_3+m_4+m_5\}/5$			
Metric Value	USEABILITY Metric Value (C_2) = $\{A_1+A_2+A_3+A_4\}/4$			

TEMPLATE 3

CRITERION: **ADAPTABILITY** (C_3)

ATTRIBUTES	METRICS	YES	NO	VALUE
Portability (A_1)	1. Can the process be applied across system configurations (mainframe - mini - micro)? (m_1)			
	2. Can the process be applied across processing methods (batch/interactive)? (m_2)			
	3. Can the process be applied across environments (Isolated, terminal room, distributed network)? (m_3)			
	4. Can the process be applied across all phases of the system life cycle? (m_4)			
	PORTABILITY Attribute Value $(A_1)=\{m_1+m_2+m_3+m_4\}/4$			
Modifiability (A_2)	1. Are input values retained by the method in their original form? (m_1)			
	2. Are calculations segmented by identifiable partitions? (m_2)			
	MODIFIABILITY Attribute Value $(A_2)=\{m_1+m_2\}/2$			
Metric Value	ADAPTABILITY Metric Value (C_3) = $\{A_1+A_2\}/2$			

TEMPLATE 4

CRITERION: FEASIBILITY (C_4)

ATTRIBUTES	METRICS	YES	NO	VALUE
Availability (A_1)	1. Is all expert opinion required for the method internal to the organization? (m_1)			
	2. Is all data required for the method internal to the organization? (m_2)			
	3. Is the data collection convenient at the scope desired? (m_3)			
	AVAILABILITY Attribute Value (A_1)=$\{m_1+m_2+m_3\}/3$			
Practicality (A_2)	1. Does the method allow input data to be in a variety of forms? (m_1)			
	2. Can the process be performed with the available staff? (m_2)			
	3. Can the process be accomplished in the time available? (m_3)			
	4. Can the precision required by the process be obtained economically? (m_4)			
	PRACTICALITY Attribute Value (A_2)=$\{m_1+m_2+m_3+m_4\}/4$			
Scope (A_3)	1. Is the amount of detail in the process user selectable? (m_1)			
	2. Does the method bound the detail at the level desired? (m_2)			
	3. Does the method analyze all hardware aspects of the system? (m_3)			
	4. Does the method analyze all software aspects of the system? (m_4)			
	5. Does the method analyze all data aspects of the system? (m_5)			
	6. Does the method analyze the procedural aspects of the system? (m_6)			
	7. Does the method analyze all personnel aspects of the system? (m_7)			
	8. Does the method analyze all communications aspects of the system? (m_8)			
	9. Does the method analyze the environment that the system resides in? (m_9)			
	SCOPE Attribute Value (A_3)=$\{m_1+m_2+m_3+m_4+m_5+m_6+m_7+m_8+m_9\}/9$			
Metric Value	FEASIBILITY Metric Value (C_4) = $\{A_1+A_2+A_3\}/3$)			

TEMPLATE 5

CRITERION: COMPLETENESS (C_5)

ATTRIBUTES	METRICS	YES	NO	VALUE
Scope (A_1)	1. Is the amount of detail in the process user selectable? (m_1)			
	2. Does the method bound the detail at the level desired? (m_2)			
	3. Does the method analyze all hardware aspects of the system? (m_3)			
	4. Does the method analyze all software aspects of the system? (m_4)			
	5. Does the method analyze all data aspects of the system? (m_5)			
	6. Does the method analyze the procedural aspects of the system? (m_6)			
	7. Does the method analyze all personnel aspects of the system? (m_7)			
	8. Does the method analyze all communications aspects of the system? (m_8)			
	9. Does the method analyze the environment that the system resides in? (m_9)			
	SCOPE Attribute Value $(A_1)=\{m_1+m_2+m_3+m_4+m_5+m_6+m_7+m_8+m_9\}/9$			
Elements (A_2)	1. Does the method comprehensively consider assets? (m_1)			
	2. Does the method comprehensively consider threat agents? (m_2)			
	3. Does the method comprehensively consider threat events? (m_3)			
	4. Does the method comprehensively consider safeguards? (m_4)			
	5. Does the method comprehensively consider vulnerabilities? (m_5)			
	6. Does the method consider outcomes? (m_6)			
	ELEMENTS Attribute Value $(A_2)=\{m_1+m_2+m_3+m_4+m_5+m_6\}/6$			
Element Attributes (A_3)	1. Are asset values considered? (m_1)			
	2. Is the potency of a threat agent considered? (m_2)			
	3. Is the undesirability of a threat event considered? (m_3)			
	4. Is safeguard effectiveness considered? (m_4)			
	5. Is the severity of an outcome considered? (m_5)			
	6. Is the likelihood (probability) of the occurrence of a threat event considered? (m_6)			
	ELEMENT ATTRIBUTES Value $(A_3)=\{m_1+m_2+m_3+m_4+m_5+m_6\}/6$			
Metric Value	COMPLETENESS Metric Value (C_5) = $\{A_1+A_2+A_3\}/3$			

TEMPLATE 6

CRITERION: **VALIDITY** (C_6)

ATTRIBUTES	METRICS	YES	NO	VALUE
Relevancy (A_1)	1. Are the results/recommendations of the process expressed in categories of solutions rather than specifics? (m_1)			
	2. Do the results relate to the significant areas of need? (m_2)			
	3. Does the process fulfill mandated requirements or regulations? (m_3)			
	RELEVANCY Attribute Value (A_1)=$\{m_1+m_2+m_3\}/3$			
Scope (A_2)	1. Is the amount of detail in the process user selectable? (m_1)			
	2. Does the method bound the detail at the level desired? (m_2)			
	3. Does the method analyze all hardware aspects of the system? (m_3)			
	4. Does the method analyze all software aspects of the system? (m_4)			
	5. Does the method analyze all data aspects of the system? (m_5)			
	6. Does the method analyze the procedural aspects of the system? (m_6)			
	7. Does the method analyze all personnel aspects of the system? (m_7)			
	8. Does the method analyze all communications aspects of the system? (m_8)			
	9. Does the method analyze the environment that the system resides in? (m_9)			
	SCOPE Attribute Value (A_2)=$\{m_1+m_2+m_3+m_4+m_5+m_6+m_7+m_8+m_9\}/9$			
Practicality (A_3)	1. Does the method allow input data to be in a variety of forms? (m_1)			
	2. Can the process be performed with the available staff? (m_2)			
	3. Can the process be accomplished in the time available? (m_3)			
	4. Can the precision required by the process be obtained economically? (m_4)			
	PRACTICALITY Attribute Value (A_3)=$\{m_1+m_2+m_3+m_4\}/4$			
Metric Value	VALIDITY Metric Value (C_6) = $\{A_1+A_2+A_3\}/3$			

TEMPLATE 7

CRITERION: **CREDIBILITY** (C_7)

ATTRIBUTES	METRICS	YES	NO	VALUE
Intuitiveness (A_1)	1. Does the method delineate the relationships between the elements? (m_1)			
	2. Does the output have a perceivable relationship with the inputs? (m_2)			
	3. Does the method analyze all hardware aspects of the system? (m_3)			
	4. Does the method analyze all software aspects of the system? (m_4)			
	5. Does the method analyze all data aspects of the system? (m_5)			
	6. Does the method analyze the procedural aspects of the system? (m_6)			
	7. Does the method analyze all personnel aspects of the system? (m_7)			
	8. Does the method analyze all communications aspects of the system? (m_8)			
	9. Does the system analyze the environment that the system resides in? (m_9)			
	INTUITIVENESS Attribute Value $(A_1)=\{m_1+m_2+m_3+m_4+m_5+m_6+m_7+m_8+m_9\}/9$			
Reliability (A_2)	1. Does the process provide a mechanism to reduce the introduction of personal bias? (m_1)			
	2. Does the method provide a mechanism that reduces the impact of uncertainty? (m_2)			
	RELIABILITY Attribute Value $(A_2)=\{m_1+m_2\}/2$			
Metric Value	CREDIBILITY Metric Value (C_7) $\{A_1+A_2\}/2$			

APPENDIX B. EVALUATION WORKSHEETS

TEMPLATE 1

CONSISTENCY					
	Reliability				
	1				
	2				
	Consistent Terminology				
	1				
	2				
	3				
	4				
USEABILITY					
	Error Handling				
	1				
	2				
	3				
	Simple				
	1				
	2				
	3				
	4				
	Ease of Use				
	1				
	2				
	3				
	4				
	Understandable				
	1				
	2				
	3				
	4				
	5				

TEMPLATE 2

ADAPTABILITY					
	Portability				
	1				
	2				
	3				
	4				
	Modifiability				
	1				
	2				
FEASIBILITY					
	Availability				
	1				
	2				
	3				
	Practicality				
	1				
	2				
	3				
	4				
	Scope				
	1				
	2				
	3				
	4				
	5				
	6				
	7				
	8				
	9				

TEMPLATE 3

COMPLETENESS					
	Scope				
	1				
	2				
	3				
	4				
	5				
	6				
	7				
	8				
	9				
	Elements				
	1				
	2				
	3				
	4				
	5				
	6				
	Element Attributes				
	1				
	2				
	3				
	4				
	5				
	6				

TEMPLATE 4

VALIDITY					
	Relevancy				
	1				
	2				
	3				
	Scope				
	1				
	2				
	3				
	4				
	5				
	6				
	7				
	8				
	9				
	Practicality				
	1				
	2				
	3				
	4				
CREDIBILITY					
	Intuitiveness				
	1				
	2				
	3				
	4				
	5				
	6				
	7				
	8				
	9				
	Reliability				
	1				
	2				

TEMPLATE 5

CONSISTENCY				
Reliability				
Consistent Terms				
USEABILITY				
Error Handling				
Simple				
Ease of Use				
Understandable				
ADAPTABILITY				
Portable				
Modifiable				
FEASIBILITY				
Availability				
Practicality				
Scope				
COMPLETENESS				
Scope				
Elements				
Element Attributes				
VALIDITY				
Relevancy				
Scope				
Practicality				
CREDIBILITY				
Intuitiveness				
Reliability				

APPENDIX C. METHOD EVALUATION WORKSHEETS

TEMPLATE 1

	ALE	CHECKLIST	SCENARIO	QUESTIONNAIRE
CONSISTENCY	**.625**	**.750**	**.250**	**.625**
Reliability	.500	1.000	.500	.500
Consistent Terms	.750	.500	.000	.750
USEABILITY	**.579**	**.817**	**.567**	**.546**
Error Handling	.666	.666	.666	.333
Simple	.500	1.000	.250	.750
Ease of Use	.750	1.000	.750	.500
Understandable	.400	.600	.600	.600
ADAPTABILITY	**.625**	**.250**	**.750**	**.500**
Portable	.750	.000	1.000	1.000
Modifiable	.500	.500	.500	.000
FEASIBILITY	**.333**	**.546**	**.694**	**.472**
Availability	.333	.666	1.000	.333
Practicality	.000	.750	.750	.750
Scope	.666	.222	.333	.333
COMPLETENESS	**.500**	**.185**	**.722**	**.722**
Scope	.666	.222	.333	.333
Elements	.500	.333	1.000	1.000
Element Attributes	.333	.000	.833	.833
VALIDITY	**.444**	**.435**	**.583**	**.583**
Relevancy	.666	.333	.666	.666
Scope	.666	.222	.333	.333
Practicality	.000	.750	.750	.750
CREDIBILITY	**.639**	**.666**	**.472**	**.367**
Intuitiveness	.777	.333	.444	.222
Reliability	.500	1.000	.500	.500

TEMPLATE 2

		ALE	CHECKLIST	SCENARIO	QUESTIONNAIRE
CONSISTENCY		.625	.750	.250	.625
	Reliability	.500	1.000	.500	.500
	1		✓		
	2	✓	✓	✓	✓
	Consistent Terminology	.750	.500	.000	.750
	1	✓			✓
	2	✓			✓
	3	✓	✓		
	4		✓		✓
USEABILITY		.579	.817	.567	.546
	Error Handling	.666	.666	.666	.333
	1				
	2	✓	✓	✓	
	3	✓	✓	✓	✓
	Simple	.500	1.000	.250	.750
	1		✓		
	2	✓	✓		✓
	3		✓		✓
	4	✓	✓	✓	✓
	Ease of Use	.750	1.000	.750	.500
	1		✓	✓	
	2	✓	✓	✓	✓
	3	✓	✓		
	4	✓	✓	✓	✓
	Understandable	.400	.600	.600	.600
	1	✓			
	2	✓	✓	✓	✓
	3		✓	✓	✓
	4			✓	✓
	5		✓		

TEMPLATE 3

		ALE	CHECKLIST	SCENARIO	QUESTIONNAIRE
ADAPTABILITY		.625	.250	.750	.500
	Portability	.750	.000	1.000	1.000
	1	✓		✓	✓
	2	✓		✓	✓
	3	✓		✓	✓
	4			✓	✓
	Modifiability	.500	.500	.500	.000
	1		✓		
	2	✓		✓	
FEASIBILITY		.333	.546	.694	.472
	Availability	.333	.666	1.000	.333
	1		✓	✓	
	2		✓	✓	✓
	3	✓		✓	
	Practicality	.000	.750	.750	.750
	1			✓	✓
	2		✓	✓	
	3		✓		✓
	4		✓	✓	✓
	Scope	.666	.222	.333	.333
	1	✓		✓	✓
	2				
	3	✓			
	4	✓			
	5				
	6		✓	✓	✓
	7	✓			
	8	✓			
	9	✓	✓	✓	✓

TEMPLATE 4

		ALE	CHECKLIST	SCENARIO	QUESTIONNAIRE
COMPLETENESS		.500	.185	.722	.722
	Scope	.666	.222	.333	.333
	1	✓		✓	✓
	2				
	3	✓			
	4	✓			
	5				
	6		✓	✓	✓
	7	✓			
	8	✓			
	9	✓	✓	✓	✓
	Elements	.500	.333	1.000	1.000
	1	✓		✓	✓
	2			✓	✓
	3	✓	✓	✓	✓
	4			✓	✓
	5		✓	✓	✓
	6	✓		✓	✓
	Element Attributes	.333	.000	.833	.833
	1	✓		✓	✓
	2				
	3			✓	✓
	4			✓	✓
	5	✓		✓	✓
	6	✓		✓	✓

TEMPLATE 5

		ALE	CHECKLIST	SCENARIO	QUESTIONNAIRE
VALIDITY		.444	.435	.583	.583
	Relevancy	.666	.333	.666	.666
	1			✓	✓
	2	✓			
	3	✓	✓	✓	✓
	Scope	.666	.222	.333	.333
	1	✓		✓	✓
	2				
	3	✓			
	4	✓			
	5				
	6		✓	✓	✓
	7	✓			
	8	✓			
	9	✓	✓	✓	✓
	Practicality	.000	.750	.750	.750
	1			✓	✓
	2		✓	✓	
	3		✓		✓
	4		✓	✓	✓
CREDIBILITY		.639	.666	.472	.367
	Intuitiveness	.777	.333	.444	.222
	1	✓		✓	
	2	✓	✓	✓	
	3	✓			
	4	✓			
	5				
	6		✓	✓	✓
	7	✓			
	8	✓			
	9	✓	✓	✓	✓
	Reliability	.500	1.000	.500	.500
	1		✓		
	2	✓	✓	✓	✓

APPENDIX D. CASE SYSTEM

A. PHYSICAL ENVIRONMENT

The subject system (patterned after a Naval Postgraduate School microcomputer network laboratory) is located on federal property whose access is controlled at night and over weekends. The laboratory is located on the second floor of a building that remains unlocked 24 hours a day. All elements of the system are located in a single room to the building, secured by a single solid core door with a metal frame. Access to the lab is controlled by a mechanical cipher lock.

The room size measures approximately 30 feet by 25 feet, the walls are standard gypsum plaster construction. The floor is carpeted wall to wall. There are no windows in the room, ventilation is provided by a single air conditioning duct sourced in the building's central air conditioner. Illumination is provided by double florescent ceiling lights.

Electrical power for the laboratory is provided by standard building power. Sufficient circuits are provided to avoid overloads. Surge and spike protection is provided for each suite of computers.

Each computer and its monitor are anchored physically to a desk. The detachable keyboards are not anchored.

B. EQUIPMENT

1. Hardware

The microcomputer network laboratory is comprised of three individual networks (3Comm Ethernet, IBM Token Ring, and Appletalk). There are 26 personal computers and seven Macintosh Pluses all of which are located inside the lab. Two IBM clones and one Macintosh are used as servers for their respective networks. The servers are write protected and keyboard locked. Six of the PC suites are equipped with 1200 baud modems each attached to a standard telephone line. Dial-in to the network is not available. There are three printers, two IBM Graphics Printers and one Apple Lazerwriter II. Total value of the hardware associated with this laboratory (including the network boards and network operating system) is approximately $62,000.

2. Software

The laboratory is used for instructional purposes during classroom hours, however, the lab is also made available as a computing facility when classes are not in session. To make the lab more useful to the student population in general, several software packages are provided to furnish basic applications. Consequently, network versions of word processing, spreadsheet, and database applications have been placed on the server. Total value of the software associated with this lab is $1715.

C. LABORATORY MANAGEMENT AND PROCEDURES

The stated purpose of the laboratory is to provide instructional microcomputer facilities to support curriculum requirements and to provide word processing, database,

and spreadsheet application services to authorized students. A laboratory policy has been written and is enforced by a part-time lab manager (he is also an instructor). The cipher lock combination is changed approximately every six months, and access to the combination is loosely controlled.

Personnel that utilize the laboratory vary in microcomputer experience from complete novices to experts with years of e' rience.

On-line virus checks are requested of the user prior to conducting any processing to avoid system infection by user provided software, and individual suite fixed disks are purged of unauthorized files on a monthly basis.

APPENDIX E. TOOL EVALUATION WORKSHEETS

TEMPLATE 1

	RiskPac	LAVA	RiskCalc	NPS
CONSISTENCY	**.375**	**.875**	**.625**	**.750**
Reliability	.000	1.000	.500	.500
Consistent Terms	.750	.750	.750	1.000
USEABILITY	**.666**	**.838**	**.738**	**.817**
Error Handling	.666	1.000	1.000	.666
Simple	1.000	.750	1.000	1.000
Ease of Use	1.000	1.000	.750	1.000
Understandable	.000	.600	.200	.600
ADAPTABILITY	**1.000**	**.625**	**.625**	**750**
Portable	1.000	.750	.750	1.000
Modifiable	1.000	.500	.500	.500
FEASIBILITY	**.583**	**.722**	**.759**	**.926**
Availability	.000	1.000	1.000	1.000
Practicality	.750	.500	.500	1.000
Scope	1.000	.666	.778	.778
COMPLETENESS	**1.000**	**.833**	**.559**	**.482**
Scope	1.000	.666	.778	.778
Elements	1.000	1.000	.500	.500
Element Attributes	1.000	.833	.400	.167
VALIDITY	**.694**	**.500**	**.759**	**.784**
Relevancy	.333	.333	1.000	.333
Scope	1.000	.666	.778	.778
Practicality	.750	.500	.500	1.000
CREDIBILITY	**.389**	**.945**	**.639**	**.583**
Intuitiveness	.778	.889	.778	.666
Reliability	.000	1.000	.500	.500

TEMPLATE 2

		RiskPac	LAVA	RiskCalc	NPS
CONSISTENCY		.375	.875	.625	.750
	Reliability	.000	1.000	.500	.500
	1		✓		✓
	2		✓	✓	
	Consistent Terminology	.750	.750	.750	1.000
	1	✓	✓		✓
	2	✓	✓	✓	✓
	3	✓	✓	✓	✓
	4			✓	✓
USEABILITY		.666	.838	.738	.817
	Error Handling	.666	1.000	1.000	.666
	1		✓	✓	
	2	✓	✓	✓	✓
	3	✓	✓	✓	✓
	Simple	1.000	.750	1.000	1.000
	1	✓	✓	✓	✓
	2	✓	✓	✓	✓
	3	✓	✓	✓	✓
	4	✓		✓	✓
	Ease of Use	1.000	1.000	.750	1.000
	1	✓	✓	✓	✓
	2	✓	✓		✓
	3	✓	✓	✓	✓
	4	✓	✓	✓	✓
	Understandable	.000	.600	.200	.600
	1		✓		✓
	2				✓
	3		✓		✓
	4				
	5		✓	✓	

TEMPLATE 3

		RiskPac	LAVA	RiskCalc	NPS
ADAPTABILITY		**1.000**	**.625**	**.625**	**.750**
	Portability	1.000	.750	.750	1.000
	1	✓	✓	✓	✓
	2	✓	✓	✓	✓
	3	✓	✓	✓	✓
	4	✓			✓
	Modifiability	1.000	.500	.500	.500
	1	✓		✓	✓
	2	✓	✓		
FEASIBILITY		.583	.722	.759	.926
	Availability	.000	1.000	1.000	1.000
	1		✓	✓	✓
	2		✓	✓	✓
	3		✓	✓	✓
	Practicality	.750	.500	.500	1.000
	1				✓
	2	✓	✓	✓	✓
	3	✓			✓
	4	✓	✓	✓	✓
	Scope	1.000	.666	.778	.778
	1	✓			
	2	✓			✓
	3	✓	✓	✓	✓
	4	✓	✓	✓	✓
	5	✓	✓	✓	✓
	6	✓		✓	✓
	7	✓	✓	✓	✓
	8	✓	✓	✓	
	9	✓	✓	✓	✓

TEMPLATE 4

		RiskPac	LAVA	RiskCalc	NPS
COMPLETENESS		1.000	.833	.559	.482
	Scope	1.000	.666	.778	.778
	1	✓			
	2	✓			✓
	3	✓	✓	✓	✓
	4	✓	✓	✓	✓
	5	✓	✓	✓	✓
	6	✓		✓	✓
	7	✓	✓	✓	✓
	8	✓	✓	✓	
	9	✓	✓	✓	✓
	Elements	1.000	1.000	.500	.500
	1	✓	✓	✓	
	2	✓	✓		
	3	✓	✓		✓
	4	✓	✓	✓	✓
	5	✓	✓	✓	✓
	6	✓	✓		
	Element Attributes	1.000	.833	.400	.167
	1	✓	✓	✓	
	2	✓	✓		
	3	✓	✓		
	4	✓	✓		✓
	5	✓	✓	✓	
	6	✓			

TEMPLATE 5

		RiskPac	LAVA	RiskCalc	NPS
VALIDITY		.694	.500	.759	.704
	Relevancy	.333	.333	1.000	.333
	1		✓	✓	
	2			✓	
	3	✓		✓	✓
	Scope	1.000	.666	.778	.778
	1	✓			
	2	✓			✓
	3	✓	✓	✓	✓
	4	✓	✓	✓	✓
	5	✓	✓	✓	✓
	6	✓		✓	✓
	7	✓	✓	✓	✓
	8	✓	✓	✓	
	9	✓	✓	✓	✓
	Practicality	.750	.500	.500	1.000
	1				✓
	2	✓	✓	✓	✓
	3	✓			✓
	4	✓	✓	✓	✓
CREDIBILITY		.389	.945	.639	.583
	Intuitiveness	.778	.889	.778	.666
	1		✓		
	2		✓		
	3	✓	✓	✓	✓
	4	✓	✓	✓	✓
	5	✓	✓	✓	✓
	6	✓		✓	✓
	7	✓	✓	✓	✓
	8	✓	✓	✓	
	9	✓	✓	✓	✓
	Reliability	.000	1.000	.500	.500
	1		✓		✓
	2		✓	✓	

LIST OF REFERENCES

Barclay, S., Brown, R.V., Kelly, C.W., Peterson, C.R., Phillips, L.D., and Selvidge, J., *Handbook for Decision Analysis*, Defense Advanced Research Projects Agency, 1977.

Boehm, B.W., Brown, J.R., Kaspar, H., Lipow, M., MacLeod, G.J., and Merrit, M.J., *Characteristics of Software Quality*, North-Holland, 1978.

Browne, P.S., "A Descriptive Risk Management Framework," *Proceedings of the 1989 Computer Security Risk Management Model Builder's Workshop, June 20-22, 1989, Ottawa, Canada*, National Computer Security Center, 1989.

Bruce, W.S., Kandel, A., and Avni, E., "The Modeling of Computer Security Systems Using Fuzzy Set Theory," *Engineering Risk and Hazard Assessment, Vol II*, Kandel and Avni (Eds.), CRC Press, 1988.

Cooper, J.A., *Computer and Communications Security: Strategies for the 1990's*, McGraw-Hill, 1989.

Fischhoff, B., "Debiasing.", pp.422-444, *Judgement Under Uncertainty: Heuristic and Biases*, Kahneman, Slovic, and Tverskey (Eds.), Cambridge University, 1982.

Fischhoff, B., Lichtenstein, S., Slovic, P., Derby, S.L., and Keeney, R.L., *Acceptable Risk*, Cambridge University, 1981.

Fowler, F.J., *Survey Research Methods*, SAGE Publications, Inc., 1984.

Gardner, G.T. and Gould, L.C., "Public Perceptions of the Risks and Benefits of Technology," *Risk Analysis, Official Journal of the Society for Risk Analysis*, Vol 9, No 2, pp.225-242, June, 1989.

Gilbert, I., Personal Interview with Authors, National Institute of Standards and Technology, Gaithersburg, Maryland, 11 August, 1989.

Glaseman, S., Turn, R., and Gaines, R.S., "Problem Areas in Computer Security Assessment," *1977 National Computer Conference, AFIPS Conference Proceedings*, American Federation of Information Processing Societies, Inc., 1977.

Hoffman, L.J., "Risk Analysis and Computer Security: Bridging the Cultural Gaps.", *Proceedings of the 1986 National Computer Security Conference*, National Computer Security Center, 1986.

Hutt, A.E., Bosworth, S., and Hoyt, D.B., *Computer Security Handbook*, MacMillan Publishing Company, 1988.

Johnson, R.E., "Playing The ODDS", *Infosystems*, April, 1987.

Katzke, S.W., "A Government Perspective on Risk Management of Automated Information Systems," *Proceedings of the 1988 Computer Security Risk Management Model Builder's Workshop, May 24-26 1988, Denver Colorado*, National Bureau of Standards, 1988.

Kahnemann, D., Slovic, P., and Tverskey, A., *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge University Press, 1982.

Krauss, L.I., *SAFE: Security Audit and Field Evaluation for Computer Facilities and Information Systems*, Firebrand, Krauss & Company, 1972.

Mayerfeld, H.T., "Framework for Risk Management: A Synthesis of the Working Group Reports From the First Computer Security Risk Management Model Builder's Workshop.", *Proceedings of the 1989 Computer Security Risk Management Model Builder's Workshop, June 20-22, 1989, Ottawa, Canada*, National Computer Security Center, 1989.

McCall, J.A., Richards, P.K., and Walters, G.F., *Factors in Software Quality (3 Volumes) (RADC-TR-77-369)*, General Electric Company for Rome Air Development Center, 1977.

Merkhofer, M., "Comparitive Analysis of Formal Decision-Making Approaches," *Risk Evaluation and Management*, Covello, Menkes, and Mumpower (Eds.), Plenum Press, 1986.

Merkhofer, M., *Decision Science and Social Risk Management*, D. Reidel Publishing, 1987.

National Bureau of Standards, *Guideline for Automatic Data Processing Risk Analysis (Federal Information Processing Standards Publications (FIPS PUB) 65)*, U.S. Department of Commerce, 1979.

National Computer Security Center, *Department of Defense Trusted Computer System Evaluation Criteria (Orange Book)*, Government Printing Office, December 1985.

Patton, M.Q., *Qualitative Evaluation Methods*, Sage Publications, 1980.

Pinsky, S., Personal Interview with Authors, National Institute of Standards and Technology, Gaithersburg, Maryland, 11 August, 1989.

Pollack, S.M., "Innovation in Systems Science Education: The Modelling Studio," *Education in Systems Science*, Bayraktar, Muller-Merbach, Roberts, and Simpson (Eds.), Halsted Press, 1979.

Quade, E.S., *Analysis for Public Decisions*, American Elsevier Publishing, 1975.

Sackman, H., The RAND Corporation, *Delphi Critique*, Lexington Books, 1974.

Schmucker, K.J., *Fuzzy Sets, Natural Language Computations and Risk Analysis*, Computer Science Press, 1984.

Shrader-Frechette, K.S., *Risk Analysis and Scientific Method*, D. Reidel Publishing, 1985.

Slovic, P., Fischhoff, B., and Lichtenstein, S., "Facts versus Fears: Understanding Perceived Risk", *Judgement Under Uncertainty: Heuristics and Biases*, Kahneman, Slovic, and Tverskey (Eds.), Cambridge University Press, 1982.

Stevens, J.A. and Weiner, R.E., "A Structured Approach to Risk Management: An Innovative Concept," *Proceedings of the 12th National Computer Security Conference*, National Institute of Standards and Technology, 1989.

Von Neumann, J., and Morgenstern, O., *Theory of Games and Economic Behavior*, Princeton University Press, 1947.

Wan, F.Y.M., *Mathematical Models and Their Analysis*, Harper and Row Publishers, 1989.

Zadeh, L.A. "Foreword," *Combining Fuzzy Imprecision With Probabilistic Uncertainty in Decision Making: Lecture Notes in Economics and Mathematical Systems*, Kacprzyk and Fedrizzi (Eds.), Springer-Verlag, 1988.

INITIAL DISTRIBUTION LIST

1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22304-6145	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93943-5002	2
3.	Computer Technology Curricular Office Code 37 Naval Postgraduate School Monterey, California, 93943	2
4.	Lance J. Hoffman Department of Electrical Engineering and Computer Science George Washington University Washington, D.C. 20052	2
5.	Magdi Kamel Code 54KA Department of Administrative Science Naval Postgraduate School Monterey, California 93943	2
6.	Major William M. Garrabrants 5243 Castle Hills Drive San Diego, California 92109	2
7.	Major Alfred W. Ellis III 99 Second Street Newport, Rhode Island 02840	2
8.	Commandant of the Marine Corps Code TE 06 Headquarters, U.S. Marine Corps Washington, D.C. 20380-0001	1
9.	Commandant of the Marine Corps Code CCIS Headquarters, U.S. Marine Corps Washington, D.C. 20380-0001	1

10. David Hsiao — 1
Code 52HQ
Department of Computer Science
Naval Postgraduate School
Monterey, California 93943

11. Sylvan Pinsky — 1
National Computer Security Center
9800 Savage Road
Fort George G. Meade, Maryland 20755-6000

12. Stuart Katzke — 1
National Institute of Standards and Technology
Gaithersburg, Maryland 20899

13. Richard Pethia — 1
Software Engineering Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213-3890