AFIT/GEE/ENV/95D-02

HOW LOCATION
AND AN
ACTION STANDARD ON A RISK LADDER
AFFECT PERCEIVED RISK

THESIS

Lonny P. Baker, Captain, USAF

AFIT/GEE/ENV/95D-02

AFIT/GEE/ENV/95D-02

HOW LOCATION
AND AN
ACTION STANDARD ON A RISK LADDER
AFFECT PERCEIVED RISK

THESIS

Presented to the Faculty of the Graduate School of Engineering

of the Air Force Institute of Technology

Air University

in Partial Fullment of the Requirements for the Degree of

Master of Science in Engineering and Environmental Management

Lonny P. Baker, B.S.

Captain, USAF

December 1995

AFIT/GEE/ENV/95D-02

HOW LOCATION
AND AN
ACTION STANDARD ON A RISK LADDER
AFFECT PERCEIVED RISK

THESIS

Lonny P. Baker, B.S.

Captain, USAF

Presented to the Faculty of the Graduate School of Engineering

of the Air Force Institute of Technology

Air University

in Partial Fullment of the Requirements for the Degree of

Master of Science in Engineering and Environmental Management

Member

Member

Chairman

Acknowledgements

First, thanks go to God, through whom all things are possible. With his help, even an old, nonengineering guy like me who hasn't seen the inside of a classroom in ten years can make it through Professor Reynolds statistics classes and the rest of this program.

I would also like to thank the instructors and students for allowing me to use their classes for the experiment. Thanks also go out to my thesis advisor, Dr. Kim Campbell, for her invaluable help and for her insights into the world of research. In addition, I would like to thank Professor Dan Reynolds, LtC Steve Lofgren, and Dr. David Christensen for their help and support of this effort.

Finally, I would like to express my gratitude to my lovely wife, Julie, for her patience, love, and support during the past eighteen months! Her understanding and the support of my family helped me to complete this program.

Lonny Baker

TABLE OF CONTENTS

List of Figures

List of Tables

AFIT/ENV/EN/95D-02

Abstract

Efforts to explain risk magnitudes often rely on a risk ladder in which exposure levels and associated risk estimates are arranged with low levels at the bottom of the ladder and high levels at the top. An experiment was conducted to test the hypotheses that perceived threat varies with the location of a subject's assigned level on the risk ladder, and that perceived threat varies with the presence of an action standard on the risk ladder. Air Force Institute of Technology Professional Continuing Education students were asked to assume a particular level of a hypothetical hazard in their residences, to read a brochure explaining the risks, and then to complete a questionnaire. The study found that subjects with an assigned risk level three-quarters of the way up the ladder perceived higher levels of risk than subjects with an assigned risk level one-quarter of the way up the ladder. The study also found that the presence of an action standard on a ladder may significantly affect perceived risk. Subjects who received a risk ladder without an action standard perceived a higher level of risk than subjects who received a risk ladder with an action standard placed above assigned risk level. These findings suggest that risk response can be shaped by effectively presented data alone.

HOW LOCATION
AND AN
ACTION STANDARD ON A RISK LADDER
AFFECT PERCEIVED RISK

I. Introduction

General Issue

In a democratic society, a major element of environmental risk management is communication about risk. In fact, it can be argued that risk communication is crucial to the success of risk management. There are many tools of science, engineering, and statistics which experts use to analyze risk-related information and to estimate and evaluate the probability and magnitude of risk. However, due to their technical nature, it is often difficult to convey information to the public about the magnitudes of risks to which they are exposed. As many discouraged policy-makers have discovered, the public often ignores information designed to alert them to significant and remediable risks, and thus fail to take appropriate action; yet this same public may insist on remedial action with respect to other risks that are too improbable or too irremediable to merit the attention they receive (Weinstein et al., 1991).

The primary goal of risk communication is to inform the participants in decisions about risk. However, risk communication is not a simple task. It involves multiple messages from many sources, and because these messages contain difficult and complex ideas, there is no simple way of making risk communication easy. To help understand this, it would be useful to present a definition of risk communication:

> Risk communication is any purposeful exchange of information about risks. It can formally be defined as the process of conveying or transmitting information among interested parties about: levels of health or environmental risks; the significance or meaning of health or environmental risks; and decisions, actions, or policies aimed at managing or controlling health or environmental risks. (Cohrssen and Covello, 1989:99)

How a communicator presents risk information is crucial to its interpretation by the public. Risk communicators are constantly seeking strategies that will tend to alert those who are being exposed to relatively high levels of the hazard in question, while reassuring those whose exposure is relatively low. The need to become more skilled at explaining risks is grounded in public health and similar concerns. Health and/or lives may be at stake when an agency tries to warn people about serious risks. On the other hand, "[w]hen people persist in worrying disproportionately about minor risks, the costs range from unnecessary anxiety to misused environmental protection dollars, from public policy gridlock to reduced agency credibility" (Weinstein et al., 1991:101).

Thus, it is very important to try and understand what influences the public's perception of risk. This understanding can then be used to develop more effective formats for the presentation of risk information to the public.

Specific Problem

The Department of Defense (DoD) does not want to place the public at risk. Likewise, in this era of shrinking budgets, the DoD does not want to spend more money than is necessary to remediate environmental hazards. For environmental cleanup alone, the DoD has spent in excess of $6.5 *billion* since 1984. In 1994 the Air Force portion of environmental cleanup costs was $559 *million* (Thal, 1994; Raymond, 1995). It should, therefore, be a goal of the DoD and the Air Force to effectively communicate risk to the public. It is neither wise to place members of the public at risk, nor is it good practice to spend the public's money unwisely.

There are also legal reasons why the DoD should be concerned with communicating risk effectively. Communication of risk of hazardous environment to the public is *required* by a number of state and federal statutes, regulations, and court decisions. Table 1 lists the applicability of federal laws and regulations to risk communication.

Table 1

Applicability of Laws and Regulations to Risk Communications (Lundgren, 1994:27)

Type of Risk Communication	CERCLA	EPCRA	NRDA	NEPA	OSHA	RCRA
Health communication					X	
Industrial hygiene					X	
Worker notification					X	
Hazardous waste	X					X
Solid waste						X
Environmental issues	X		X	X		X
Emergency planning		X				
Actual crisis			X			

CERCLA - Comprehensive Environmental Response, Compensation, and Liability Act
EPCRA - Emergency Planning and Community Right-to-Know Act
NRDA - National Resource Damage Assessment
NEPA - National Environmental Policy Act
OSHA - Occupational Safety and Health Act
RCRA - Resource Conservation and Recovery Act

Failing to understand and follow the laws and regulations can have serious repercussions. For instance, an organization may be sued for failing to follow due process. This has happened to a number of federal agencies because they took a law or even their own implementing regulations less seriously than did the public. Also, the agency in charge of implementing the regulation may shut down operations or levy a heavy fine if an organization is not in compliance (Lundgren, 1994).

It is important to realize that risk communication is an iterative process with many inputs and variables. Risk is communicated to the public through many media, and communication efforts are often frustrating for both communicator and recipient: "[t]his is largely due to the fact that risk messages necessarily compress technical information, which can lead to misunderstanding, confusion, and distrust" (NRC, 1989:3). These problems and widespread dissatisfaction have focused attention on risk communication.

Realizing that there are problems with risk communication, an important question to ask is where to focus efforts in improving the process? One aspect of risk communication which is receiving attention is that of risk comparisons. Lundgren states that "[a]n aspect of presenting risk information that concerns most of us who communicate about risk is how to *compare* risks" (Lundgren, 1994:58; emphasis added). Many experts, including the National Research Council (NRC), agree and argue that risk comparisons are important because in many cases they help put difficult and complex information into a familiar perspective which the public can more easily understand (Brown, 1985; Covello, 1989; NRC, 1989; Weinstein et al., 1989). Unlike scientists who are used to dealing with quantitative data and, therefore, have a frame of reference for numbers, most of the public does not have a context for this type of data (Brown, 1985). By providing careful comparisons, the NRC asserts that people can better understand this information. "Comparing different risks can help people comprehend the uncommon magnitudes involved and understand the level, or magnitude, of risk associated with a particular hazard" (NRC, 1989:96).

Risk communicators are looking for better formats which present risk comparisons both accurately and in a way that will make sense to people with no technical training. One tool used for comparing risks and explaining risk magnitudes is the *risk ladder*. Figure 1 and Figure 2 are examples of common types of risk ladders. Risk ladders are a type of graphic presentation format used for risk comparison in which a range of probabilities is presented for a single class of risks. This technique is very helpful for understanding the magnitude of a particular risk. Although there are many types of risk ladders with varying scales and hazard comparisons, most formats use a vertical exposure ladder to display different exposure levels and information about those levels. For example, in

Figure 1 a list of activities are arranged vertically, with the most risky (smoking) at the top of the graphic and the least risky (lightning) at the bottom of the graphic. Exposure levels and associated risk estimates are usually arrayed with low levels at the bottom and high levels at the top. Figure 2 is another example of a risk ladder which can be found in the Environmental Protection Agency's (EPA) A Citizen's Guide to Radon.

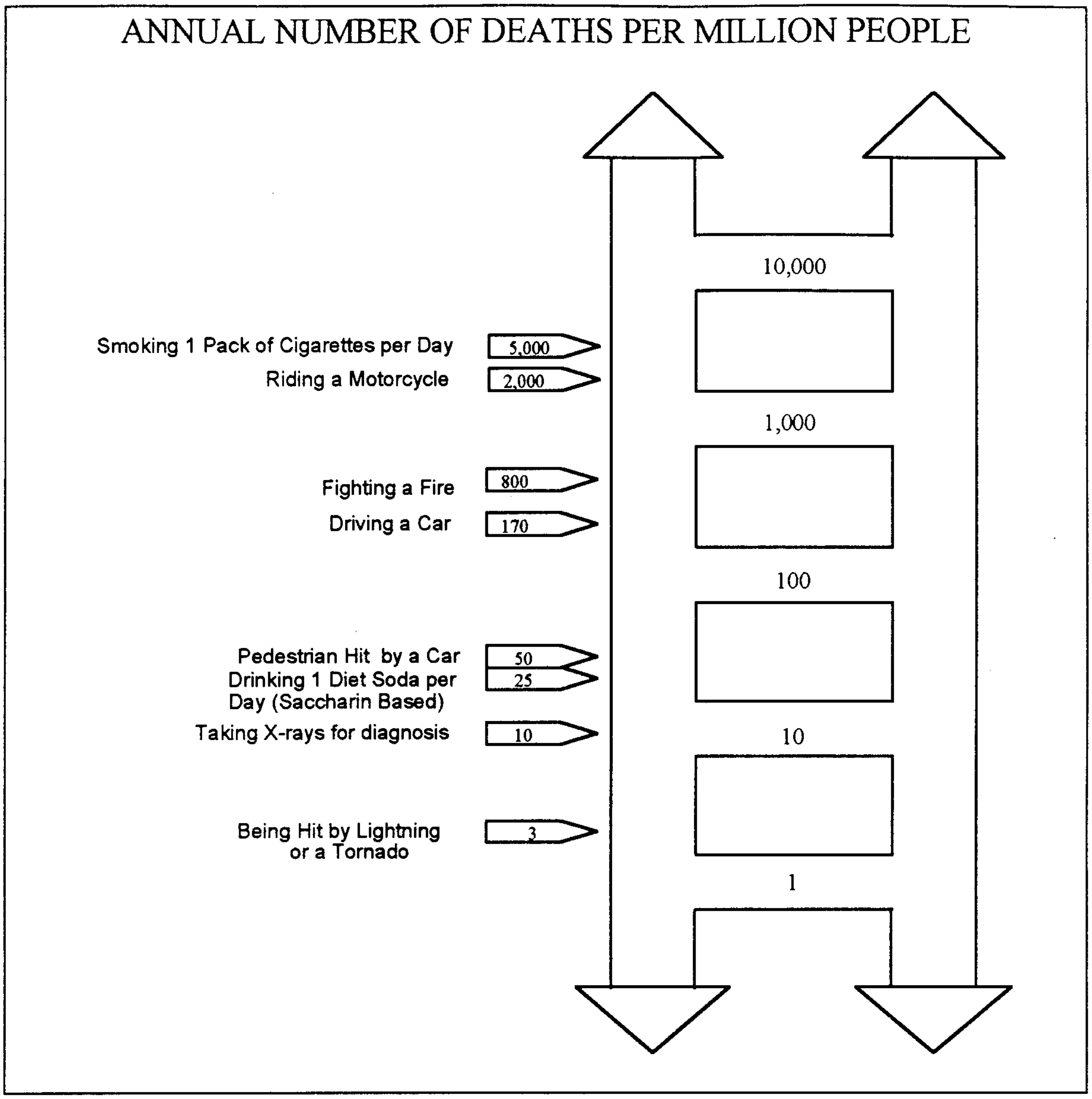


Figure 1. Example of a Risk Ladder
(Schultz et al., 1986 as cited in Covello et al., 1988)

RADON RISK IF YOU SMOKE

Radon Level	*If 1,000 people who smoked were exposed to this level over a lifetime . . .*	*The risk of cancer from radon exposure compares to . . .*	*WHAT TO DO: Stop Smoking and . . .*
20 pCi/L	About 135 people could get lung cancer	<-- 100 times the risk of drowning	Fix your home
10 pCi/L	About 71 people could get lung cancer	<-- 100 times the risk of dying in a home fire	Fix your home
8 pCi/L	About 71 people could get lung cancer		Fix your home
4 pCi/L	About 71 people could get lung cancer	<-- 100 times the risk of dying in an airplane crash	Fix your home
			Consider fixing between 2 and 4 pCi/L
2 pCi/L	About 71 people could get lung cancer	<-- 2 times the risk of dying in a car crash	
1.3 pCi/L	About 71 people could get lung cancer	(Average indoor radon level)	(Reducing radon levels below 2 pCi/L is difficult)
0.4 pCi/L	About 71 people could get lung cancer	(Average outdoor radon level)	

Note: If you are a former smoker, your risk may be lower.

Figure 2. Example of a Risk Ladder
(adapted from EPA, 1992:12)

While risk ladders are used for explaining risk magnitudes, not all the attributes of risk ladders have been empirically tested. Therefore, it is difficult to determine with certainty how people react to them. In fact, only one study which specifically tested risk ladders could be found.

It is important to note that the use of risk comparisons, whether in a risk ladder format or not, is not without problems. There are many factors other than risk magnitudes that affect risk perception which have been identified (e.g., Slovic, et al., 1985; Covello et al., 1988; NRC, 1989). However, "[t]here are hardly any data to support claims that one approach works better than another" [Weinstein et al., 1989:11). This study attempts to help fill that gap in our knowledge.

The objective of this thesis is to contribute to the research on risk ladders as a presentation format for explaining risk magnitudes to the public.

Research Questions

The following research questions will be addressed in this thesis:

1. Does the location of assigned risk level on a risk ladder affect perceived risk?
2. Does the presence of an action standard on a risk ladder affect perceived of risk?
3. Do location on a risk ladder and the presence of an action standard interact to affect perceived risk?

Scope

Researchers have studied and identified many factors other than risk magnitude that affect how the public perceives particular risks. However, "[m]uch less research has attempted to determine how to explain the magnitudes of risks, and thus improve the correlation between risk and response" (Weinstein et al., 1989:1). In particular, very limited research has been accomplished in testing the effectiveness of presentation formats for risk magnitudes.

There is a wide array of presentation formats available for use by the risk communicator. Furthermore, for each format available, there are innumerable attributes which can be studied. Of the many attributes of risk ladders, only two will be studied: location on the risk ladder and the presence of an action standard on a risk ladder. An experiment will be conducted to determine if location of assigned risk on a risk ladder affects perceived risk. Additionally, if it is found that location on a risk ladder significantly affects perceived risk, the experiment will also determine if the presence of an action standard helps to counteract this effect.

Overview

Chapter II discusses the use of risk ladders in risk communication and the relevant research which has been accomplished in this area. Chapter III explains and justifies the experimental design, pertinent concepts, construction of the experimental instrument, the procedures used for administering the experiment, and the statistical analysis used to analyze the experimental results. Chapter IV discusses the statistical analysis techniques used to test for significant differences in the data. Finally, Chapter V discusses the conclusions of the tested hypotheses and provides recommendations for further research.

II. Literature Review

Introduction

This chapter is divided into four sections. The first section discusses the need for empirical research on the attributes of risk ladders. The second section reviews the relevant research on communicating risk magnitudes. The third section summarizes the relevant findings from the EPA study from which this research was patterned. Finally, the fourth section provides concluding remarks.

The Need for Research

"Comparing different risks can help people comprehend the uncommon magnitudes involved and understand the level, or magnitude, of risk associated with a particular hazard" (NRC, 1989:96). Several experts in risk communication agree with this statement. Covello believes that because of the strengths of risk comparisons, a critical need exists to better understand factors that influence the effectiveness of risk comparisons (Covello, 1988). Furthermore, he states that empirical research is needed that addresses a diverse set of questions regarding risk comparisons, including: "What means of display are most effective for risk comparisons?" and "How can basic risk assessment terms, such as 'parts per billion,' be more effectively presented and explained through comparisons?" Sandman et al. state that "[a]lthough investigators have identified many factors other than risk magnitudes that seem to influence how the public responds to particular risks, much less research has sought the best ways of explaining risk magnitudes" (Sandman et al, 1994:35).

In 1989 the National Research Council (NRC) released a report entitled *Improving Risk Communication*. This committee recognized that as a major element in risk management, risk communication was a major problem (NRC, 1989:ix). The goal of the committee was to improve the risk communication process. In reviewing the risk communication process, the committee determined that

> Risk messages commonly convey quantitative information that is unfamiliar and difficult to comprehend. These magnitudes and risk estimates are not easily understood without benchmarks or points of reference, and providing careful comparisons can help people understand this information. (NRC, 1989:96)

The committee also believed that a technique which can help people understand the magnitudes of risk is the risk ladder (NRC, 1989:96). However, they found that not all attributes of risk ladders have been empirically tested, and that "[e]ach practical use of risk comparison should be carefully pretested if possible" (NRC, 1989:97).

Previous Research

A search of the existing scientific literature was conducted for information about communicating risk magnitudes. Books and journals from various disciplines, including communication, psychology, and public health were the focus of attention. It was hoped that sources would be found which would provide information about effective techniques for communicating risk through printed materials. However, while there has been research conducted which identifies factors that influence risk perceptions and judgments, very few studies have examined methods for explaining risk magnitudes (e.g., Slovic, et al., 1985; Covello et al., 1988; NRC, 1989). Additionally, only a single two-phase study could be found which specifically tested various presentation formats for explaining risk magnitudes (Weinstein et al., 1989; Weinstein et al., 1991).

After finding little information about communicating risk magnitudes in printed form, the search turned to the field of graphic communication in hopes of finding research relevant to the risk ladder. While many research efforts have focused on improving the appearance of graphs (American National Standards Institute, 1979; Enrick, 1980; Szoka et al., 1991), others have focused on improving the analytical usefulness of graphs (Barber and Dunn, 1992; Christensen and Larkin, 1992; Tan and Benbasat, 1993). For instance, Enrick describes desired characteristics of appearance such as margins, text, lettering, scales, titling, symbols, and abbreviations (Enrick, 1980). On the other hand, other sources deal with more analytical aspects of graphics. For example, Christensen and Larkin identified nine criteria for high integrity graphics, and reported that ". . . graphs which violate the criteria can mislead decision makers" (Christensen and Larkin, 1992:130). The nine criteria are as follows:

(1) The graph should agree with the data (known as Tufte's *lie factor*).

(2) Follow normal temporal and sign conventions.

(3) In area charts, the area with the least variability should be on the bottom.

(4) Labels should be correct.

(5) The number of dimensions in the graph should not exceed the number of dimensions in the data.

(6) Avoid unusual scaling.

(7) The scale range should be close to the data range.

(8) Avoid arbitrary changes to grid proportions.

(9) Beware of omitted data. (Christensen and Larkin, 1992:131-144)

Risk ladders, however, do not present information in what most would consider a common graphical format. The *Encyclopaedia Britannica* defines graph as a "... pictorial representation of statistical data or of a functional relationship between variables" (Encyclopaedia Britannica, 1994:429). Most graphs employ two axes, in which the vertical axis represents a group of dependent variables, and the horizontal axis represents a group of independent variables. Data points are usually plotted on such a grid and then connected with lines, as illustrated in Figure 3.

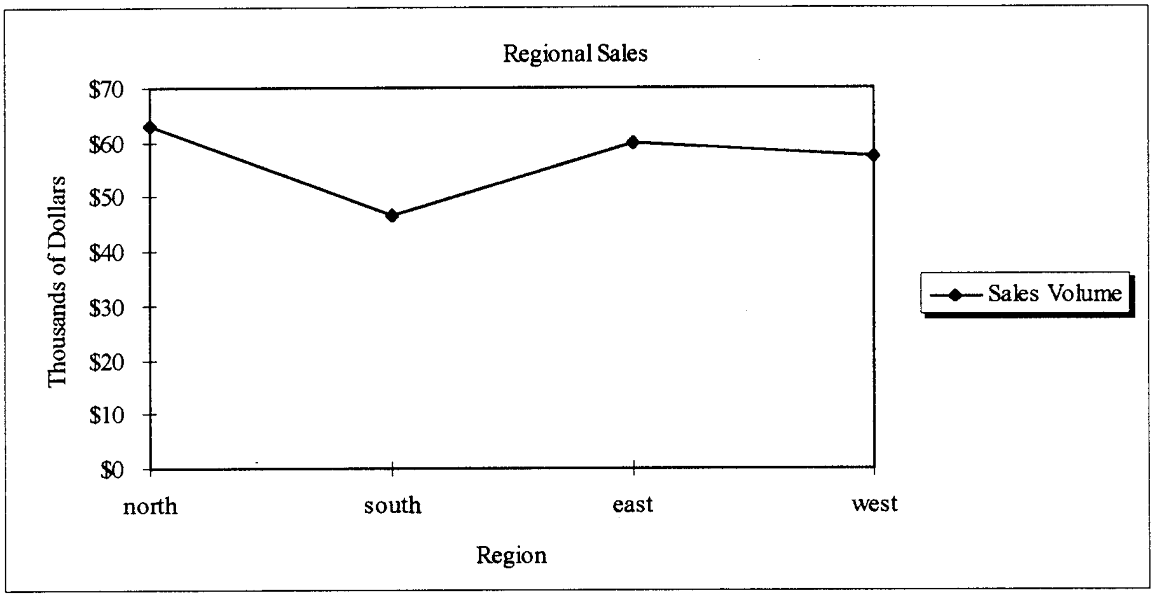


Figure 3. Example of a Common Graph

Other examples of common graphs include pie charts, bar graphs, and surface graphs. Most graphs are a picture which illustrate the functional relationship between independent and dependent variables as changes or differences in area, height, or slope.

Unlike most common types of graphs, the risk ladder is not a picture which illustrates the functional relationship between the independent and dependent variables. Rather, it usually lists the independent variable data (exposure to a hazard) in a vertical fashion and simply lists the corresponding dependent variable data (risk of health effects) alongside in a table-like fashion. The risk ladder also allows for additional information to be included, such as comparisons to other risks and advice, to be displayed more easily.

EPA Research on Risk Ladders

The only research on risk ladders which could be located was a two-phase study conducted as part of a Cooperative Agreement between the Environmental Communication Research Program of Cook College, Rutgers University, and the Office of Policy, Planning and Evaluation of the U.S. Environmental Protection Agency. The study, entitled *Communicating Effectively About Risk Magnitudes*, examined a variety of risk presentation formats and tested their success in communicating about two different hazards, geological radon and asbestos. The results of the first phase were published in August of 1989, and the results of the second phase were published in September of 1991.

Phase One (Weinstein et al., 1989).

Phase one of the research evaluated seven presentation formats. Of the seven formats tested, five displayed exposure levels in the form of a risk ladder, one used histogram bars, and one contained no visual representation of exposure levels. Four of the formats included an action standard below which mitigation was not recommended, and above which mitigation was recommended. Based on the results, the researchers

> . . . believe that in the absence of a ladder, subjects have no way of telling what types of results are unusual, and they react therefore as though the finding of any amount of hazardous substance were a serious problem. (Weinstein et al., 1989:49)

This reinforces an NRC statement that ". . . [risk] magnitudes and risk estimates are not easily understood without benchmarks or points of reference . . ." (NRC, 1989:96).

The phase one research found that the presence of an action standard increased the likelihood that subjects' action intentions matched the action recommendations. In other words, those subjects with exposure levels below the action standard had perceptions of risk which were lower than those with exposure levels above the action standard.

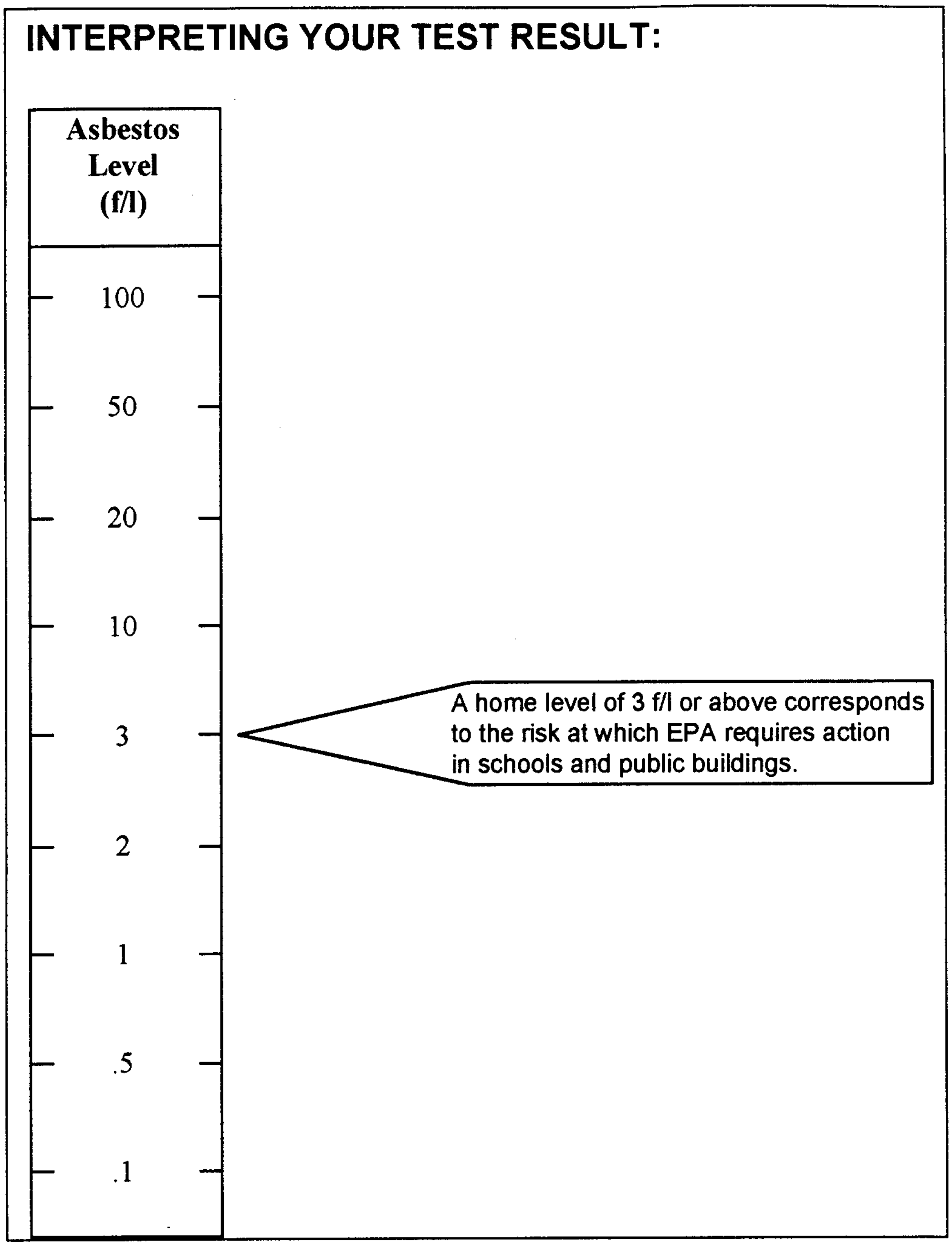
Also of significance in the phase one study was the development of the 'locational hypothesis.' Although location on the page wasn't specifically tested, based on observation, the researchers felt that within-hazard threat perceptions and mitigation intentions were primarily a product of the placement of assigned risk on the page, rather than an appreciation of the magnitude of the risk. Subsequently, this hypothesis was specifically tested in phase two of the study.

Phase Two (Weinstein et al., 1991).

Based primarily on the results of the first phase, phase two tested location on the page, units of exposure magnitude, simultaneous presentation of two hazards, and other hypotheses. The various hypotheses were tested through three different experiments.

In phase one, the standard-only format had produced the most risk averse response of any of the formats tested. The researchers hypothesized "that the strongly risk-averse response to the standard-only format might result from the fact that it was the only condition that did not contain a ladder or other visual representation of exposure levels" (Weinstein et al., 1991:27). Therefore, experiment I of phase two was designed to test the hypothesis that simply adding a risk ladder to the standard-only format would affect perceived risk.

The formats for experiment I consisted of a standard-only format and a ladder containing exposure levels (low levels of exposure at the bottom and high at the top) with an action standard located at the midway point. The standard-plus-ladder format is shown in Figure 4. Four hypothetical test result levels were used (0.8, 2.5, 3.5, and 24.0 fibers/liter), with subjects' assigned risk levels either slightly above, well above, slightly below, or well below the action standard. The results showed that the effect on perceived threat was slightly stronger for those subjects exposed to the standard-only format than for those subjects exposed to the standard-plus-ladder format.

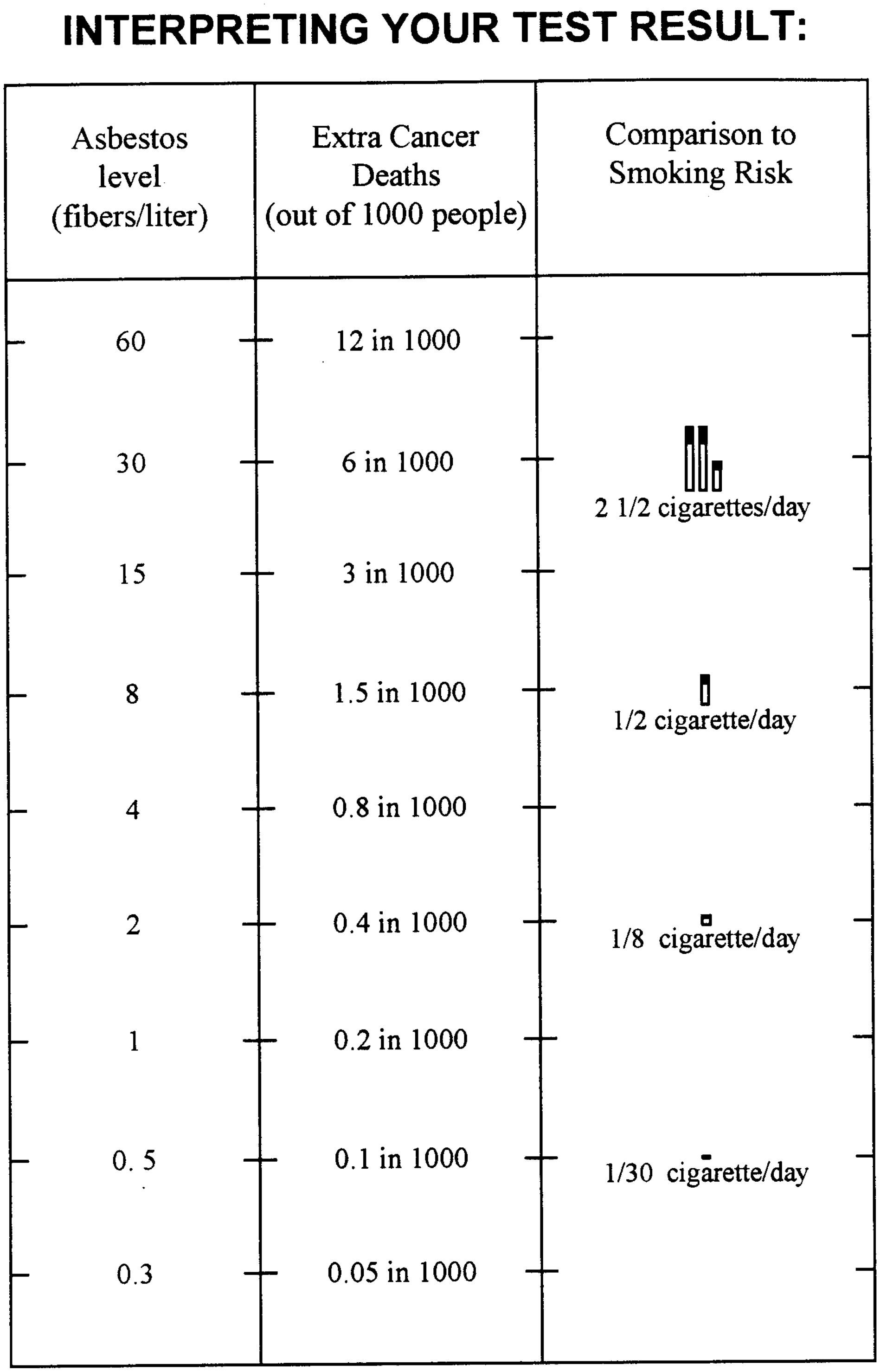

INTERPRETING YOUR TEST RESULT:

A home level of 3 f/l or above corresponds to the risk at which EPA requires action in schools and public buildings.

Figure 4. Example Format from EPA Study, Phase Two, Experiment I
(Weinstein et al., 1991:A-5).

Based on observations in phase one, the 'locational hypothesis' was specifically tested in experiments II and III of phase two. Both experiments used basically the same formats, and used geological radon and asbestos as the hazards. An example of the basic format is shown in Figure 5.

The locational hypothesis was tested by displacing the risk ladder. In other words, the same hypothetical test result with the same risk information was located either one-quarter of the way up the ladder or three-quarters of the way up the ladder. Both experiments found a significant locational effect on perceived threat. In other words, subjects with assigned risk levels located high on the ladder perceived their risk as being higher than subjects with the same assigned risk level but located low on the ladder, despite the same level of actual risk.



INTERPRETING YOUR TEST RESULT:

Asbestos level (fibers/liter)	Extra Cancer Deaths (out of 1000 people)	Comparison to Smoking Risk
60	12 in 1000	
30	6 in 1000	2 1/2 cigarettes/day
15	3 in 1000	
8	1.5 in 1000	1/2 cigarette/day
4	0.8 in 1000	
2	0.4 in 1000	1/8 cigarette/day
1	0.2 in 1000	
0. 5	0.1 in 1000	1/30 cigarette/day
0.3	0.05 in 1000	

Figure 5. Example Format from EPA Study, Phase Two, Experiment III
(Weinstein et al., 1991:144)

Summary

Various presentation formats are used by risk communicators to help communicate risk magnitudes to the public. While a substantial amount of research has identified many of the factors other than risk magnitudes that seem to determine how the public responds to particular risks, very little research has investigated the best ways of explaining risk magnitudes. In fact, other than the EPA funded study mentioned here, there has been relatively little empirical research to support claims that one presentation format works better than another. The risk communicator is faced with the task of choosing a presentation format which produces the appropriate level of response to a certain level of risk. In choosing a presentation format, its effect on an individual's perceptions should be considered.

The goal of this research is to contribute to our knowledge of one particular presentation format: the risk ladder. The EPA has initiated the study of this format and this thesis will continue that effort in two ways. First, the test of the locational hypothesis in phase two, experiment III will be replicated. An experiment will be conducted to see if perception of risk varies with location on a risk ladder. Second, the locational hypothesis will be tested in conjunction with the action standard hypothesis. In the EPA study the action standard was only located at the midpoint of the risk ladder. This experiment will test the locational hypothesis by varying the location of the action standard as well. The next chapter discusses the details of the methods used in this thesis.

III. Methodology

Introduction

This chapter describes the methodology used to research the hypotheses concerning whether or not location on a risk ladder and the presence of an action standard on a risk ladder affect perceived risk. There are five sections in this chapter: section one discusses the subjects used in the research; section two describes the overall experimental design; the development of the experimental instrument is covered in section three; the fourth section explains how the data were collected; and, section five discusses the characteristics of sound measurement.

Subjects

Seventy-eight subjects from the AFIT Professional Continuing Education (PCE) program voluntarily participated. All were employees of the United States Air Force attending professional continuing education logistics courses at AFIT. The subjects came from varying backgrounds but the majority are currently working in the logistics, acquisition, maintenance, quality assurance, or related technical career fields. Sixteen subjects were female and sixty-two were male. Sixteen were enlisted, varying in rank from E-6 to E-8, with one E-3. Thirty-three subjects were officers, ranging in rank from O-1 to O-5, and twenty-nine subjects were civilian employees. The average age was 36.8 years while the median was 37 years. Experience level in their current career field ranged from 1 to 26 years, with an average of 11.3 years. The subjects were highly educated with the average education level being at least completion of a 4-year college degree, while the median education level was completion of some graduate study. Eleven of the seventy-eight subjects were smokers.

AFIT PCE students were chosen primarily because they were an extremely accessible source. Therefore, it must be noted that the subjects were not selected *randomly* but, rather, *conveniently*. "Although random selection enables a researcher to generalize results to a population, one may need to settle for a convenience sample because an entire group of individuals (e.g., a classroom, an organization, a family unit) is available to participate in the study" (Creswell, 1994:127).

There are experts who feel that some types of nonrandom samples, such as convenience samples, are of little use to the researcher. It is felt that these subjects do not represent any defined population, and the results from studies using such samples are useless in ascertaining public sentiment (Chadwick, 1984). However, other researchers feel that the results of an experiment which relies on a convenience sample may be generalized beyond the single experiment (Parsons, 1974; Keppel, 1991). The use of a convenience sample presumes that there is no reason for any one section of the population to be different than any other chunk with respect to the characteristic under investigation. In fact,

> . . . past research in a number of laboratories with subjects chosen from different sources (for example, different breeding stocks, different suppliers of laboratory animals, and human subjects from different schools in different sections of the country) have shown that these differences are relatively unimportant in the study of various phenomena. Knowing this, an investigator working in this field may feel safe in generalizing the results beyond the single experiment. (Keppel, 1991: 17-18)

Keppel (1991) maintains that there is a distinction between a *statistical* generalization, which depends on random sampling, and a *nonstatistical* generalization, which depends on knowledge of a particular research area. Cornfield and Tukey (1956) and Edgington (1966) conclude that the extension of a set of findings to a broader class of subjects is dictated primarily by what is known in a particular field of research about the appropriateness of certain generalizations and the length of these generalizations. "The availability of this information will depend on the state of development of the research area and the extent to which extrapolations beyond the particular subjects tested have been successful in the past" (Keppel, 1991:18). Based on the discussion above, the results of the research reported here can with good conscience be generalized to the career fields aforementioned, and possibly to the entire Air Force or beyond.

Experimental design

The EPA study used a single-factor between-subjects design with four levels of the treatment condition (Weinstein et al., 1991), while this experiment used a completely randomized, 2 x 3, full factorial design as illustrated in Figure 3.

		Location	
		high	low
Format	no action standard	1. . 13	14. . 26
	action standard above	27. . 39	40. . 52
	action standard below	53. . 65	66. . 78

Figure 3. Experimental Design

The independent variables were ***location*** and ***format***. ***Location*** had two levels: high on the page and low on the page. ***Format*** had three levels: no action standard, action standard above assigned risk level, and action standard below assigned risk level. The dependent variable was ***perceived risk***. There were thirteen subjects per treatment for a total of 78 subjects.

The factorial, or multifactor, design was chosen because it has several advantages over single factor designs (Neter et al., 1990; Keppel, 1991). First, factorial designs are more efficient because they permit the manipulation of more than one independent variable in the same experiment. Additionally, factorial designs provide more information about the main and interaction effects because "[t]he arrangement of the treatment conditions is such that information can be obtained about the influence of each of the independent variables considered *separately* and about how the variables *combine* to influence behavior" (Keppel, 1991:19). Finally, factorial designs can strengthen the validity of the findings because they permit movement beyond a single-dimensional view to a richer and more revealing multidimensional view (Keppel, 1991).

In addition to being fully factorial, the experiment was also a completely randomized or between-subjects design. This design is characterized by the fact that subjects are randomly assigned to, and serve in only *one* of, the different treatment conditions. “Any differences in behavior observed between any one treatment condition and the others are based on the differences between independent groups of subjects” (Keppel, p. 19). Although it is not necessary, equal numbers of subjects were assigned to each treatment group. The advantages of between-subjects designs are that they are

> . . . simpler to understand conceptually, are easier to design and to analyze, and are relatively free from restrictive statistical assumptions. The main disadvantages are the large number of subjects required for even a modest experiment and a relative lack of sensitivity in detecting treatment effects when they are present. (Keppel, 1991:19)

Experimental Instrument

The experimental instrument used in this research (see Appendix A) was a modification of the instrument used in the EPA phase two study. The EPA experimental instrument included a short brochure which described a hazard. Their two-page brochure was developed from information contained in various government publications and the hazards used were either radon or asbestos. Each subject was to imagine that his home had been tested for a hazard and then use the hypothetical test result, along with a risk ladder, to answer some questions. The first two pages, with basic information about the hazard, were constant across all treatments. The last page of the brochure was the format being tested. Attached to the end of the brochure were some questions designed to measure the dependent variables. These were followed by a feedback questionnaire. This study modified the asbestos brochure as described below.

Hazard Selection.

It was felt that some subjects might have preconceived notions about asbestos and that these feelings would bias the results. Therefore, a fictitious hazard named *fibronite* was used in place of asbestos. The wording of the original brochure was also modified in hopes that this would prevent some of the bias that might be associated with asbestos. In the post-test questionnaire nearly all of the subjects indicated that they recognized the similarities between asbestos and *fibronite*. However, due to the wording of the question, it is unclear whether or not the subjects recognized the similarities before or after reading the question. Regardless of the fact that most said they recognized the similarities, the significance of the experimental results indicates that there probably was no strong bias associated with the imaginary hazard *fibronite*.

Risk Ladders.

The six risk ladders used in this experiment (see Appendix B) were nearly identical to those in the EPA study. However, in addition to giving the hypothetical test result in the brochure, the test result was also indicated on the ladder. In this way, any errors in reading the ladder were prevented.

Action standards were also included on four of the six risk ladders in order to see if the location on the page had the same effect on threat perception if the location was accompanied by an action standard either above or below the assigned risk level. The

1991 EPA study did not include action standards in the locational experiment. In the earlier study conducted in 1989, the EPA research only included an action standard at the midway point on the page. Assigned risk levels were then varied in location either slightly below/above or far below/above the action standard.

In this study, the location of assigned risk level was either one-quarter of the way up the ladder or three-quarters of the way up the ladder. For each location an action standard was placed either just below the test result or just above the test result.

Questionnaire.

The questionnaire used in this research (see Appendix A) is a subset of the items in the questionnaire from the 1991 EPA study. Four questions were taken verbatim from the EPA questionnaire and used as the risk-related response measures. A fifth item was a composite index of perceived threat created from the four questions. The questions were designed to measure responses related to:

> (1) *Perceived likelihood.* A 7-point scale assessed subjects' perception of the likelihood of harmful effects from their hypothetical test results. The scale values ranged from 1 = no chance to 7 = certain.
>
> (2) *Perceived Seriousness.* A 6-point scale assessed perceived seriousness, with values ranging from 1 = no risk to 6 = very serious risk.
>
> (3) *Concern.* Concern was measured using a 5-point scale ranging from 1 = not at all concerned to 5 = extremely concerned.
>
> (4) *Fear.* Fear was measured using a 5-point scale ranging from 1 = not at all frightened to 5 = extremely frightened.
>
> (5) *Composite Index of Perceived Threat.* The four items pertaining to perceived likelihood, perceived seriousness, concern, and fear were added together to form a composite index of perceived threat which ranged in value from 4 to 23. (Weinstein et al., 1991:35-36)

The composite index was developed to provide a more sensitive and reliable response measure (Weinstein et al., 1991). These four variables were highly inter-correlated, with inter-item correlations ranging from 0.57 to 0.76 and a coefficient alpha of 0.84. This suggests they tap the same general dimension (Weinstein et al., 1991:36).

The composite index used in the 1991 EPA study is more commonly known as a *summated rating scale*. Across the social sciences, summated rating scales are used quite often.

> A political scientist may pose several items to survey respondents about their 'trust in government,' adding up scores to form an index for each. A sociologist might ask a sample of workers to evaluate their 'subjective social class' in a battery of questions, summing responses into one measure. In each example, the goal is development of an individual rating on some attitude, value, or opinion. (Spector, 1991:v)

There are four characteristics that make a scale a summated rating scale:

(1) A scale must contain multiple items - the use of *summated* in the name implies that multiple items will be combined or summed.

(2) Each individual item must measure something that has an underlying, quantitative measurement continuum - it measures a property of something that can vary quantitatively rather than qualitatively. An attitude, for example, can vary from being very favorable to being very unfavorable.

(3) Each item has no 'right' answer - this makes the summated rating scale different from a multiple-choice test and, therefore, cannot be used to test for knowledge or ability.

(4) Each item is a statement, and respondents are asked to give ratings about each statement - it involves asking subjects to indicate which of several response choices best reflects their response to the item. Most summated rating scales offer between four and seven response choices. (Spector,1991:1)

The summated rating scale format is used for several reasons. First, it can produce scales that have both good reliability and validity. Second, it is relatively cheap and easy to develop. Finally, a well-devised scale is usually quick and easy for respondents to complete (Spector, 1991:2). Each of these reasons is discussed in more detail below.

The advantages of a summated rating scale over typical yes-or-no and true-false type questions concern reliability, precision, and scope (Spector, 1991:4). Single items are unreliable because a person may answer 'yes' today and 'no' tomorrow. Additionally, single items are imprecise because measurement is restricted to only two levels. People can only be placed into two groups, and it is impossible to distinguish among people in each group. "Finally, many measured characteristics are broad in scope and not easily assessed with a single question. Some issues are complex, and several items will be necessary to assess them" (Spector, 1991:4). "A good summated rating scale is both *reliable* and *valid*" (Spector, 1991:6).

Reliability should be considered in two ways; test-retest reliability and internal-consistency reliability (Spector, 1991:6). Test-retest reliability means that a scale yields consistent measurement over time. Test-retest reliability cannot be assessed for the scale used in this study, because it was not possible to conduct more than one test. Additionally, the 1991 EPA study did not report any assessments of test-retest reliability. Internal-consistency reliability means that multiple items, designed to measure the same construct, will intercorrelate with one another (Spector, 1991:6). Internal consistency, as measured by inter-item correlation and coefficient alpha, was reported in the 1991 EPA study, and will be assessed in this study using the same method.

> Internal consistency is a measurable property of items that implies that they measure the same construct. It reflects the extent to which items intercorrelate with one another. . . . Internal consistency among a set of items suggests that they share common variance or that they are indicators of the same underlying construct. The nature of that construct or constructs is certainly open to question. (Spector, 1991:30)

As stated, the internal consistency of a summated rating scale is measured by calculating the inter-item correlation and the coefficient alpha. Inter-item correlation is determined by calculating the *item-remainder* coefficient for each item. (This statistic is also known as the *part-whole* or *item-whole* coefficient.) The item-remainder coefficient is the correlation of each item with the sum of the remaining items (Spector, 1991:30). There are three basic strategies for deciding which items to retain based on the value of the item-remainder coefficient:

> (1) If it is decided that the scale should have *m* items, then the *m* items with the largest coefficients would be chosen.
>
> (2) A criterion for the coefficient (e.g., 0.40) can be set, and all items with coefficients at least that great would be retained.
>
> (3) Both strategies can be used together. Up to *m* items can be retained, providing they have a minimum sized coefficient. (Spector, 1991:31)

This study will use the criterion method and all coefficients with a value of at least 0.50 will be retained.

Coefficient alpha, or Cronbach's alpha, is a measure of the internal consistency of a scale and should be computed for any multiple-item scale (Spector, 1991; Carmines and Zeller, 1979). A widely accepted rule of thumb is that alpha should be at least 0.70 for a scale to demonstrate internal consistency (Nunnaly, 1978).

The second requirement of a good summated rating scale is validity. Validation, that is, interpreting what the scale scores represent, is the most difficult part of scale development.

> If it is internally consistent, the scale certainly measures something, but determining the nature of that something is a complex problem. Part of the difficulty is that validation can only occur within a system of hypothesized relations between the construct of interest and other constructs. Tests for validity involve simultaneous tests of the hypotheses about constructs and hypotheses about scales. (Spector, 1991:46)

This type of validation is beyond the scope of this thesis and will not be addressed. Therefore, the assumption will be made that the scale used for this study is valid.

Feedback Questionnaire.

The subjects filled out a feedback questionnaire after completion of the experiment. No analysis using these data was conducted. The questionnaire asked standard demographic questions such as age, occupation, and education level. The feedback questionnaire can be found in Appendix A.

Data Collection

The experimental sessions were conducted over three days and included three AFIT PCE logistics classes (Appendix E contains the raw data). Each class varied in size: the first had nineteen students; the second had fifty-one students; and the third had eight. All three sessions were conducted between the hours of 11:00 p.m. and 1:00 p.m. This author served as the experimenter for all sessions. The experimenter greeted each class and then handed each subject an informed consent form which explained the purpose of the experiment, gave the instructions for completing the experiment, and informed the subjects of their rights as participants in the experiment. A copy of the informed consent form can be found in Appendix A. Each subject was also given a brochure which included the hazard information, a risk ladder, a questionnaire, and a feedback questionnaire in that order. All sections of the brochure were constant for all subjects with the exception of the risk ladder, which was the treatment condition. The time spent by each subject on the experiment varied from ten to fifteen minutes.

Subjects were randomly assigned to the treatment conditions by using a random number table. The treatments were each assigned a different number ranging from one to six (since there were six treatment conditions). Using the random number table, a treatment was chosen, assigned to a subject, then withdrawn from the pool. The rest of the treatments were assigned in this manner until the final treatment was chosen by default. This procedure is effectively sampling without replacement. The next six subjects were assigned treatments in the same manner and this procedure was continued until all subjects had been assigned a treatment.

The critical features of the random assignment are that each subject-treatment combination is equally likely to be assigned to any one of the six treatments and that the assignment of each subject is independent of that of the others. "Following this procedure, then, we guarantee that each of the treatment conditions is equally likely to be assigned to a given subject and to whatever other uncontrolled factors might be present during any period of testing" (Keppel, 1991:16).

Characteristics of Sound Measurement

At the most general level, there are two basic properties of empirical measurements. They are reliability and validity (Cook and Campbell, 1979; Carmines and Zeller, 1982). "Fundamentally, r*eliability* concerns the extent to which an experiment, test, or any measuring procedure yields the same results on repeated trials" (Carmines and Zeller, 1982:11). "*Valid(ity)* refers to the extent that (a test) measures what it purports to measure" (Carmines and Zeller, 1982:12). A good understanding of this topic is necessary for good research. However, the complete discussion on validity and reliability is rather lengthy, and for this reason was placed in Appendix C. To summarize Appendix C, this study has fulfilled the requirements for both validity and reliability.

Data Analysis

This section is divided into two parts. The first deals with the computation of Cronbach's alpha (or coefficient alpha) which was used to assess the reliability of the construct tested in this study. The second section discusses the method used to identify differences which can be attributed to the various experimental treatments. The SAS code and the corresponding output for all procedures in this study are located in Appendix F.

Assessing Reliability of the Test Instrument.

Cronbach's alpha was used to assess the reliability of the test instrument. It involves comparisons of the variance of a total scale score, or the sum of all items, with the variances of the individual items.

The formula for Cronbach's alpha is:

$$\alpha = \frac{k}{k-1} \times \frac{s_T^2 - \sum s_I^2}{s_T^2}$$

where

α = Cronbach's alpha

k = the number of items

s_T^2 = the total variance of the sum of the items

s_I^2 = the variance of an individual item

The numerator of the equation contains the difference between the total scale variance and sum of the item variances. Once the ratio of this difference is calculated, it is multiplied by a function of the number of items. Cronbach's alpha was calculated using the Statistical Analysis System (SAS, version 6.08) on a VAX Model 6000-420, serial number 0B000006.

Assessing Perception of Risk.

Neter, Wasserman, and Kutner (1990) suggest the following basic strategy for analyzing factor effects in two-factor studies:

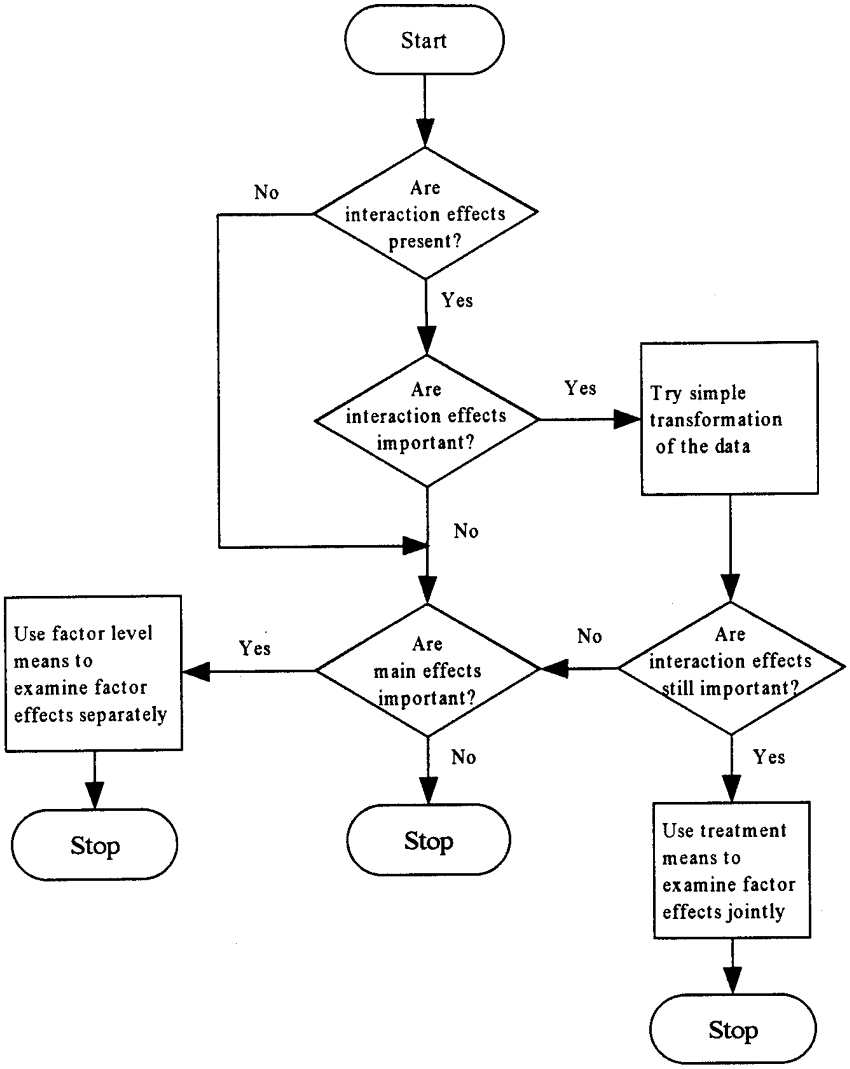


Figure 5. Strategy for Analysis of Two-Factor Studies
(Neter et al., 1990:730-731)

This strategy was followed when analyzing the factor effects in this study.

Analysis of Variance (ANOVA) procedures were performed on the dependent variable ***risk perception*** using the SAS on a VAX Model 6000-420, serial number 0B000006. Post hoc simple-effect *F* tests were performed to evaluate significant interactions. Tukey's test was used to compare means. Significance levels for all tests were controlled at $\alpha = 0.05$.

This study used a completely randomized, 2 x 3, factorial design. There were thirteen subjects per treatment for a total of seventy-eight. Each subject experienced only one treatment. The linear model underlying the analysis begins with a statement of the components contributing to any observation in the experiment (Y_{ijk}) (Keppel, 91, 218):

$$Y_{ijk} = \mu_T + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

where

μ_T = the overall mean of the population

α_i = the average treatment effect of location at level a_i $(\alpha_i = \mu_i - \mu_T)$
(i = 1, 2, where 1 = high location, 2 = low location)

β_j = the average treatment effect of format at level b_j $(\beta_j = \mu_j - \mu_T)$
(j = 1, 2, 3, where 1 = no action standard, 2 = action standard above, 3 = action standard below)

$(\alpha\beta)_{ij}$ = the interaction effect at cell a_ib_j $((\alpha\beta)_{ij} = \mu_{ij} - \mu_i - \mu_j + \mu_T)$
(the interaction of the main effects location and format)

ε_{ijk} = experimental error associated with each observation $(\varepsilon_{ijk} = Y_{ijk} - \mu_{ij})$

"Before undertaking formal inference procedures, it is desirable to evaluate the aptness of the two-factor ANOVA model" (Neter et al., 1990:705). There are three basic assumptions of ANOVA:

(1) <u>Independence of Observations</u> - independence means that each observation is in no way related to any other observations in the experiment. Independence is achieved by randomly assigning subjects to conditions and testing them individually.

(2) Normally Distributed Treatment Populations - the individual treatment populations, from which the members of each treatment group are assumed to be randomly drawn, must be normally distributed.

(3) Homogeneity of Variance - the variance of each treatment group should be equal. (Neter et al., 1990:623-625; Keppel, 1991:96-99)

The requirement of the independence of observations was satisfied by the random assignment of subjects to the treatment conditions, and by testing each subject individually. The assumption of normality was validated by the Wilk-Shapiro/Rankit Plot. A minimum value of 0.80 was required to satisfy the normality assumption, where a value of 1.0 indicates a perfectly normal distribution. The Hartley test was used to determine whether or not the variances of each treatment group were equal.

The Tukey test, which utilized the *studentized range statistic*, was used for comparison of means. The Tukey test is generally recommended for most experimental situations for the following reasons: First, this test can be used when there is interest in evaluating the significance of all possible differences between pairs of treatment means. Second, the Tukey test may be used to maintain the *Family Wise* error rate at the chosen value of α_{FW} for the entire set of pairwise comparisons. When all sample sizes are equal, the family confidence coefficient for the Tukey method is exactly $1-\alpha$. ". . . [W]hen the family confidence coefficient is $1-\alpha$, all pairwise comparisons in the family will be correct in $(1-\alpha)100$ percent of the families" (Neter et al., 1990:582).

Summary

An experiment was conducted to see if location of assigned risk level on a risk ladder and the presence of an action standard on a risk ladder affect perceived risk. The experiment used a completely randomized, 2 x 3 full factorial design. The independent variables were ***location*** and ***format***. ***Location*** had two levels; high on the page and low on the page. ***Format*** had three levels; no action standard, action standard above assigned risk level, and action standard below assigned risk level. The dependent variable was ***perceived risk*** as measured by perceived threat. There were thirteen subjects per treatment for a total of seventy-eight subjects. The experimental results were analyzed using ANOVA procedure and means were compared using Tukey's test.

IV. Findings and Analysis

Experimental Findings and Analysis

This chapter presents the data obtained from the experiment and an analysis of the results. This study used a completely randomized, 2 x 3 full factorial design. ***Location*** consisted of two levels, high and low. ***Format*** consisted of three levels, no action standard, action standard above test result, and action standard below test result. Each subject saw only one condition, and there were thirteen subjects per condition. This experimental design was used to test research questions one and two.

Appendices F-G contain the complete results of the experiment. Appendix F contains the raw data, to include the test results and demographic data, and descriptive statistics for each treatment. Appendix G contains the SAS code and the results used for the various intercorrelation, ANOVA,and Tukey tables and figures included in this chapter.

Reliability of the Test Instrument

The internal-consistency method was used to assess the reliability of the summated rating scale which was used to measure the construct ***risk (threat) perception***. The results of these calculations are displayed in Table 2.

As indicated in Table 2, the overall Cronbach coefficient alpha is 0.88. This satisfies the adopted criterion of 0.70. The item-remainder coefficient values are quite high, ranging from 0.69 to 0.78. These values are similar to those obtained in the 1991

EPA study, where inter-item correlations ranged from 0.57 to 0.76 and coefficient alpha was 0.84 for the scale (Weinstein et al., 1991:36). It should be noted that removal of none of the items from the scale will raise the coefficient alpha value higher than 0.88. Based on the Cronbach alpha value of 0.88, it was concluded that the composite index (summated rating scale) is a reliable measure of the construct, ***risk (threat) perception***.

Table 2

Cronbach's Alpha and Inter-item Correlation for the Summated Rating Scale

Cronbach's Alpha

for RAW variables: 0.876504

Deleted Item	**Item-Remainder Coefficient**	**Alpha if Item Removed**
Question 1	0.708489	0.851613
Question 2	0.777339	0.826808
Question 3	0.774684	0.829259
Question 4	0.692551	0.857610

Assumptions of the ANOVA Model

As stated in Chapter Three, the assumption of independence of observations has been met. The Hartley test was used to determine whether or not the variances of each treatment group were eaual. Based on the results of the Hartley test, it was concluded that the assumption of homogeneity of variance was not violated. Appendix D contains the complete calculations. The assumptions of independence of observations and homogeneity of variance are quite sensitive to departures from normality. The histogram and Wilk-Shapiro plot of the residuals indicate that the data meets the normality assumption (Figure 6 and Figure 7). The approximate Wilk-Shapiro value for the data is 0.9896. A Wilk-Shapiro value above 0.8 and a relatively straight 45 degree line for the Rankit Plot are indications of normality.

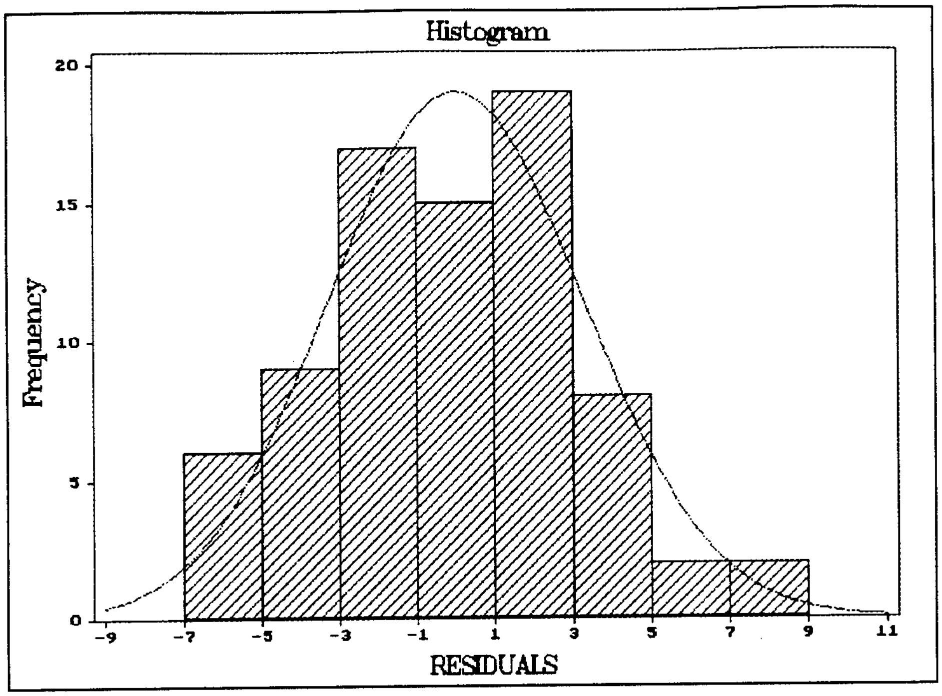


Figure 6. Histogram of Residuals

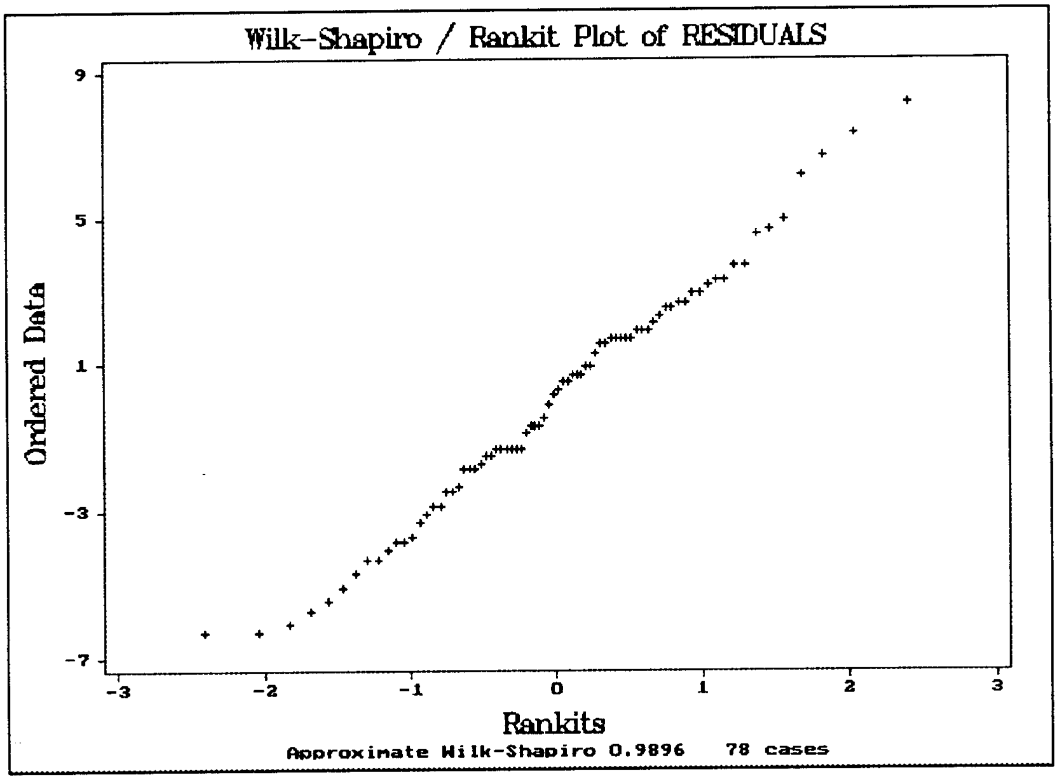


Figure 7. Wilk-Shapiro Normality Plot of Residuals

Descriptive Statistics

Table 3 shows some of the descriptive statistics for each treatment. A more complete list of descriptive statistics can be found in Appendix F. The minimum value any one subject could score was four, and the maximum value was twenty-two. An observation of Table 3 shows that no subject scored either the minimum or maximum value in any treatment, although some came close. This indicates that no subject felt either completely threatened or completely nonthreatened by the hypothetical test result.

Table 3

Descriptive Statistics by Treatment

		LOCATION		
		Low	**High**	
FORMAT	**No Action Standard**	mean: 15.31 median: 14 sd: 3.35 max: 22 min: 9	mean: 16.31 median: 17 sd: 3.33 max: 21 min: 10	mean: 15.81 std dv: 3.31
	Action Standard Above	mean: 11.85 median: 10 sd: 3.83 max: 20 min: 8	mean: 13.69 median: 13 sd: 3.59 max: 21 min: 8	mean: 12.77 std dv: 3.76
	Action Standard Below	mean: 11.46 median: 12 sd: 2.67 max: 16 min: 6	mean: 16.08 median: 17 sd: 3.45 max: 21 min: 10	mean: 13.77 std dv: 3.83
		mean: 12.87 std dv: 3.67	mean: 15.36 std dv: 3.57	mean: 14.12 std dv: 3.81

Statistical Tests for the Location and the Action Standard Hypotheses

This section describes the results of the statistical tests used to answer the research questions established in Chapter One. The research questions and their statistical hypotheses are:

(1) Does the location of assigned risk level on a risk ladder affect perceived risk?

H_o: $\mu_{low} = \mu_{high}$
(There is no difference between the location means.)

H_a: $\mu_{low} \neq \mu_{high}$
(There is a difference between the location means.)

(2) Does the presence of an action standard on a risk ladder affect perceived risk?

H_o: $\mu_{NAS} = \mu_{ASA} = \mu_{ASB}$
(There is no difference between the format means.)

H_a: $\mu_{NAS} \neq \mu_{ASA} \neq \mu_{ASB}$
(At least two of the format means differ.)

(3) Do location on a risk ladder and the presence of an action standard interact to affect perceived risk?

H_o: all $(\alpha\beta)_{ij} = 0$
(The factors location and format don't interact to affect risk perception.)

H_a: not all $(\alpha\beta)_{ij} \neq 0$
(The factors location and format do interact to affect risk perception.)

Table 4 shows the results of the two-factor ANOVA on the dependent variable ***risk perception***. The criterion level for statistical significance was set at 0.05 for all tests.

Table 4

Summary of Analysis of Variance

Source	*df*	*SS*	*MS*	*F*	*p*
MAIN EFFECTS					
A (Location)	1	120.628	120.682	10.513	0.0018*
B (Format)	2	124.692	62.346	5.434	0.0063*
INTERACTION					
AB (Location x Format)	2	46.487	23.244	2.026	0.139
Subjects/AB (error)	72	826.154	11.474		
TOTAL	77	1117.962			

* Denotes significance at $p < 0.05$

The first step is to determine if interaction effects are present. "An interaction is present when the pattern of differences associated with an independent variable changes at the different levels of the other independent variable" (Keppel, 1991:196). Figure 8 shows the plot of the interaction of the main effects. Perfectly parallel curves would indicate there are no interactions. However, the curves aren't perfectly parallel, indicating that some interaction is present. Nevertheless, the lines are closely parallel, suggesting the interaction is not significant. Results of the analysis indicate that there is no significant interaction between the main effects ***location*** and ***format***, $F(2, 72) = 2.026, p = 0.139$ (Table 2). Therefore, for research question three, the null hypothesis cannot be rejected. The factors ***location*** and ***format*** do not interact significantly to affect ***perceived risk***.

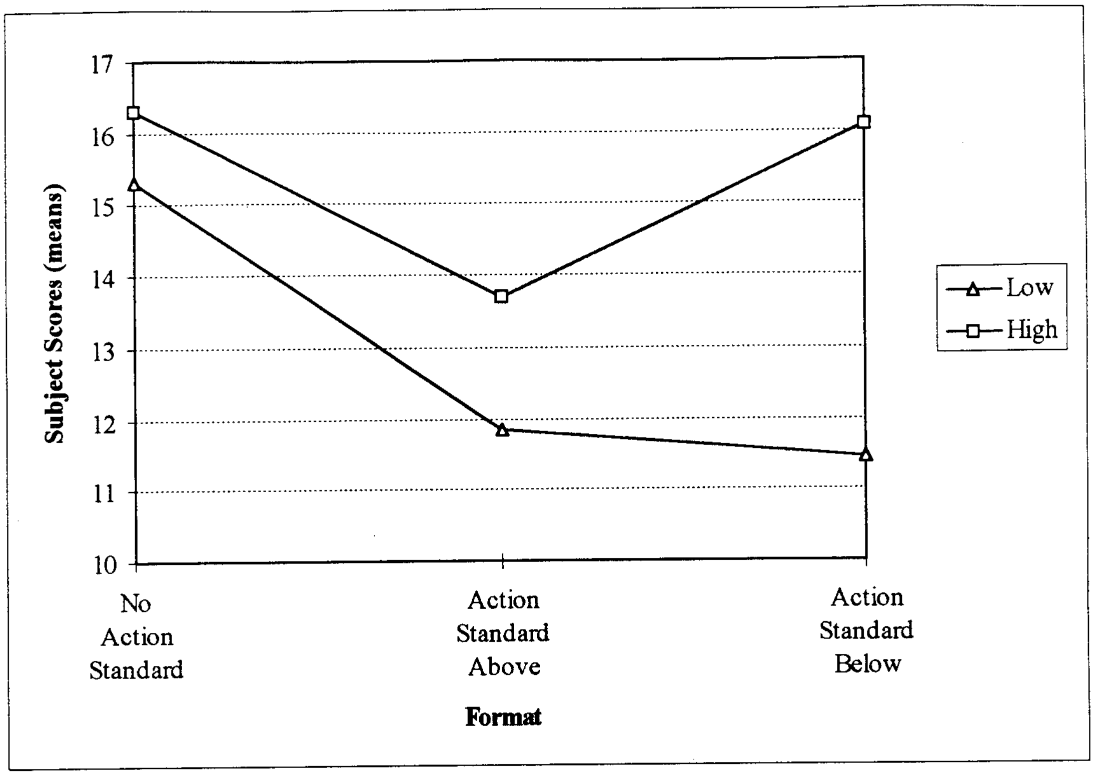


Figure 8. Plot of the Interaction of Main Effects

The next step is to test the main effects. If any of the main effects are significant, all that is necessary is to examine the factor effects separately by testing the factor level means. As shown in Table 4, the results of the analysis indicate that the main effect ***location*** is significant, $F(1, 77) = 10.51, p = 0.0018$. This is illustrated in Figure 9 (the error bars on the bar graph represent one standard deviation above and below the mean). The results indicate that a location high on the page ($M = 15.36$) increased perceptions of threat above those elicited by a location low on the page ($M = 12.87$). Therefore, the null hypothesis for research question one is rejected, resulting in the acceptance of the alternate hypothesis; location on a risk ladder significantly affects ***perceived risk***.

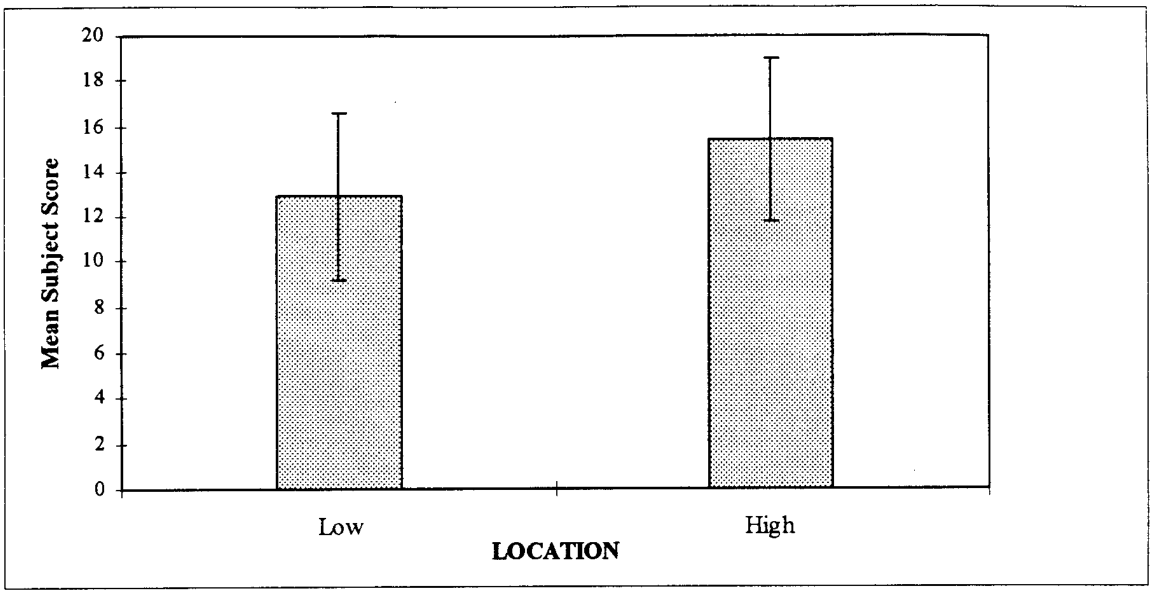


Figure 9. Mean Score as a Function of Location

Results of the analysis indicate that the main effect ***format*** is also significant, *F*(2, 77) = 5.43, *p* = 0.0063, as illustrated in Figure 10. Post hoc analysis revealed a significant difference between formats with no action standard and formats with the action standard above the assigned risk level (see Table 5).

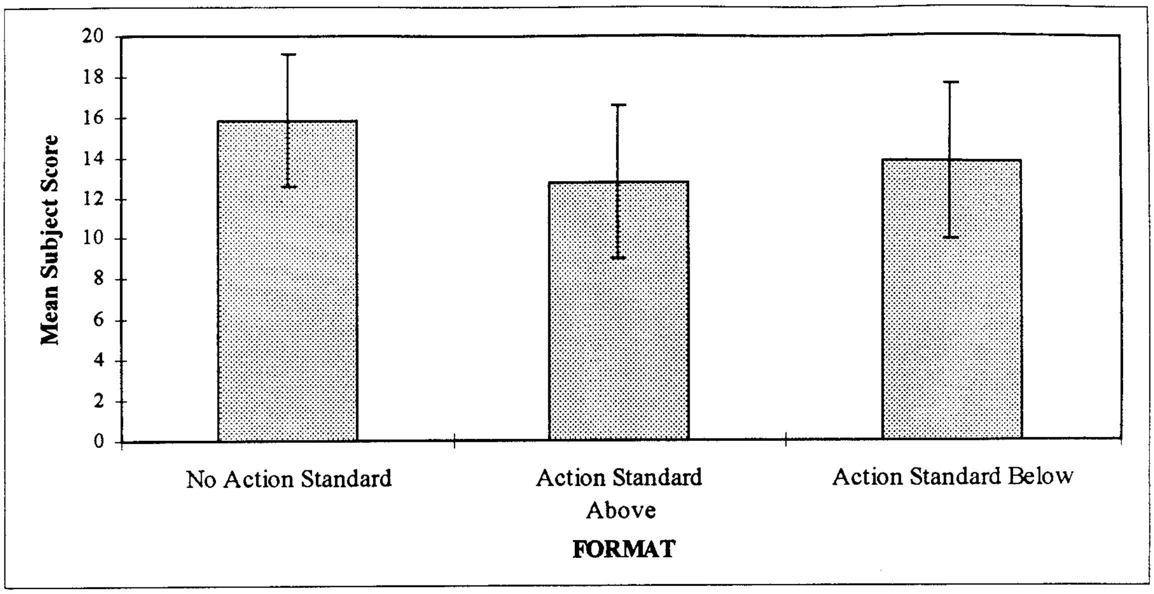


Figure 10. Mean Score as a Function of Format

Table 5

Results of Post Hoc Comparisons of Format

Minimum Significant Difference = 2.2483			
FORMAT	**Mean**	**Tukey Grouping**	
No Action Standard	15.808	A	
Action Standard Below	13.769	A	B
Action Standard Above	12.769		B

α =.05 *df* = 72 *MSE* = 11.474

That is, subjects' perception of threat was higher for the format with no action standard (*M* = 15.808) than for the format with the action standard above assigned risk level (*M* = 12.76). The format with the action standard below assigned risk level was not significantly different from either of the other levels. The null hypothesis for research question two can, therefore, be rejected. This allows acceptance of the alternate hypothesis; the presence of an action standard on a risk ladder significantly affects perceived risk.

Summary

Analysis of Variance procedures were performed to identify differences in perceived risk attributable to the various experimental treatments. This method was used to determine the main and interaction effects of the factors ***location*** and ***format*** on the dependent variable ***risk perception***. When main effects were found to be significant, Tukey's test was used to compare means. The criterion level for statistical significance was set at 0.05 for all tests. The following paragraphs summarizes the results of the analysis.

(1) Results of the analysis indicate that there is no significant interaction between the main effects ***location*** and ***format***. In other words, the main effects ***location*** and ***format*** do not significantly interact to affect ***risk perception***.

(2) Results of the analysis indicate that the main effect ***location*** is significant. Subjects with assigned levels of risk high on the ladder had higher perceptions of risk than those with assigned levels low on the ladder.

(3) Results of the analysis indicate that the main effect ***format*** is significant. Subjects without an action standard perceived a higher risk level of risk than subjects with an action standard above assigned risk level. Subjects with an action standard below assigned level of risk perceived their level of risk as neither higher nor lower than the other two formats.

V. Conclusion and Recommendations

Introduction

". . . [M]agnitudes and risk estimates are not easily understood without benchmarks or points of reference, and providing careful comparisons can help people understand this information" (NRC, 1989:96). A graphical image has the ability to communicate quantitative risk information that is unfamiliar and difficult to comprehend by providing these comparisons. Risk ladders, which present a range of probabilities for a single class of risks, are a common graphic format that are used to make risk comparisons. While many researchers believe that risk ladders are an effective tool for communicating risk magnitudes, risk ladders do have limitations and problems. Many of these problems stem from the fact that little research has empirically tested the various attributes of risk ladders.

This research examined a risk presentation format and tested its success in communicating about an imaginary hazard *fibronite*. Based on the characteristics of asbestos, *fibronite* has important properties: it confronts individuals rather than being a community-wide problem, tests can be carried out to indicate the seriousness of the risk, and individual -level mitigation is possible (Weinstein et al., 91). Various other hazards fall into this important category such as asbestos, geological radon, contamination of water by lead pipes, and contamination of home wells.

Successful communication about this type of hazard is important because it can increase the likelihood that people take actions to reduce health risks when appropriate, and can decrease the likelihood of excessive worry and unnecessary action when risk is low. When individuals aren't able to assess their own risk and make their own decisions about mitigation, different communication strategies may be required.

This study examined two attributes of risk ladders: location of risk level on a risk ladder and the location of an action standard on a risk ladder. Both attributes were found to significantly affect perceived risk.

The Effect of Location on a Risk Ladder

By displacing the risk ladder, the same hypothetical test result with the same risk information was located either one-quarter of the way up the ladder or three-quarters of the way up the ladder. Results of the analysis indicate that the test result located three-quarters of the way up the ladder elicited statistically higher perceptions of threat than the test result located one-quarter of the way up the ladder.

Although not surprising, this is an important finding. "Virtually every primer on graphical presentation stresses that graphs can be truncated, extended, or displaced in order to exaggerate or minimize the effect displayed" (Weinstein et al., 91:92). This research indicates that peoples' perceptions of risk can be altered, either arbitrarily or intentionally, by constructing a risk ladder so that their risk appears near the top or near the bottom of the ladder.

Clearly there are ethical as well as practical considerations when constructing a risk ladder. A ladder can be constructed to intentionally reduce or increase empathy or panic. To intentionally affect risk perceptions ". . . is in the judgment of some observers to leave the domain of risk education and enter that of risk propaganda" (Weinstein et al., 91:98). However, when constructing a risk ladder, each of the attributes must be decided upon. Since some of these attributes do have predictable impacts, risk communicators should be aware of those impacts and how they affect risk perception.

The Effect of an Action Standard on a Risk Ladder

In addition to determining if location on a risk ladder affects perceived risk, this study examined if the presence of an action standard on a risk ladder affects perceived risk. To some, it may seem obvious that the presence of an action standard would increase the likelihood that perceptions match recommendations. However, in the area of risk communication, what is obvious to one group of the public may not always be obvious to another. This study was also able to examine if the presence of an action standard can counteract the effects of location on a ladder.

The results of the analysis indicate that the effect of the format was significant. Further analysis revealed that a risk ladder without an action standard created higher perceptions of threat than those elicited by the ladder in which the action standard was above risk level. This is comforting, because it shows that explicit action standards significantly affect perceptions of risk.

A more disturbing result was that the ladder in which the risk level was above the action standard didn't produce a higher perception of threat than when the risk level was below the action standard. It was expected that risk levels above the action standard would have induced a stronger perception of threat. Perhaps a reason for this was that the risk levels were too close to the action standard on the ladder, and subjects perceived levels close to the standard, whether above or below it, as being equally threatening. However this reasoning is contrary to the results of the EPA Phase I study (Weinstein et al., 89). That study found that an action ". . . standard creates an artificial discontinuity in hazard response as one goes from just below the standard to just above the standard (Weinstein et al., 89:49). In fact "[m]any commentators have complained about the public's tendency to dichotomize around a standard, perceiving levels just below the standard as safe and those just above as risky" (Weinstein et al., 89:3).

Perhaps a more plausible explanation is that the locational effect caused the responses to average out. When location of risk level was low on the ladder, subjects should not have felt more threatened than when location of risk level was high, since actual risk levels were the same. However, the difference between the mean responses for the two locations for ladders with the action standard below risk level ($M_{low} = 11.46$ vs. $M_{high} = 16.08$) was the largest of any two treatments. (These values clearly show how effective location on the page can be in affecting risk perception!) As expected, the difference between the two locations wasn't as great for the ladders with the action standard above risk level ($M_{low} = 11.85$ vs. $M_{high} = 13.69$), but they averaged to within one unit of the format with the action standard below risk level. ($M_{below} = 12.77$ vs. $M_{above} = 13.77$). Thus, there was no significant statistical difference between the two formats.

Recommendations

The goal of communication about personally mitigable hazards is to increase the likelihood that people take actions to reduce health risks when appropriate, and to decrease the likelihood of unneeded action when risk levels are low. Presentation formats can be useful in this task, and risk communicators should not develop presentation formats without regard for format attributes and their known effects on perceptions. It has been empirically shown in this study that location on a risk ladder and an action standard on a risk ladder affect perceived risk.

When developing risk ladders, risk communicators should always include an action standard when possible. An explicit action standard increases the likelihood that individual perceptions and intentions match recommendations, and thus can contribute meaningfully to public health.

Additionally, when developing a risk ladder, risk communicators should be aware of the locational effect. Making intentional use of this effect can be an ethical choice, and is up to the individual practitioner. However, blindly developing a risk ladder without regard for this effect may hamper the risk communication process. Awareness of the locational effect can prevent an 'unintentional' response opposite the one which is appropriate.

Suggestions for Further Research

The possibilities for research in this area are boundless. Most of the efforts in graphics research have been accomplished only within the past decade, and a need for more empirical research remains. The use of graphics for communicating information has increased, especially considering the advances in computer technology over the past few years. The software packages which exist today have made the creation of charts and graphs almost effortless, contributing to the widespread use of such graphics.

Risk ladders have proven to be successful in communicating within-hazard risks. However, an effective format for presenting both within-hazard risks and between-hazard risks has yet to be found (Sandman et al., 1994). Another variable not yet explored is the use of graphic representations of probability or concentration data (Weinstein et al., 1991). Several EPA brochures have made considerable, and controversial, use of a matrix of faces and crosses to show mortality probabilities. However, the effect of this format on risk perception has not yet been tested. Another possibility would be to do a similar study as this one, but with formats other than risk ladders. Finally, some graphics research has indicated that people who are right-brain dominant process graphical information differently than those who are left-brain dominant (DeSanctis, 1984; MacKay and Villarreal, 1987). This suggests that gender may be a factor in the interpretation of graphics. By blocking on gender, a study could help determine if the risk perceptions of men and women are affected differently by the effects of location and an action on a risk ladder.

In conclusion, there is still much work to be done in evaluating the effect of graphical formats on risk perception. As studies continue to contribute to the understanding of graphics and their effect on human perception, more effective and more useful graphical formats will be discovered.

Bibliography

American National Standards Institute. American National Standard Time-series Chart. New York: The American Society of Mechanical Engineers, 1979.

Barber, Ellen T. And Elizabeth C. Dunn. The Effects of Iconic Presentation on Individuals. MS Thesis, AFIT/GCA/LSY/92S-1. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1992 (AD-A258785).

Brown, Rex V. Presenting Risk Management Information to Policymakers: Executive Summary. Falls Church, VA: Decision Science Consortium, July 1985

Carmines, Edward G., and Richard A. Zeller. Reliability and Validity Assessment, Series: Quantitative Applications in the Social Sciences. Beverly Hills: Sage Publications, 1982.

Chadwick, Bruce A. Research Methods. Englewood Cliffs, NJ: Prentice-Hall, 1984.

Christensen, David S. And Albert Larkin. “Criteria for High Integrity Graphics,” Journal of Managerial Issues, 4-1: 130-153 (1992).

Cohrssen, John J. And Vincent T. Covello. Risk Analysis: A Guide to Principles and Methods for Analyzing Health and Environmental Risks. Washington D.C.: U.S. Council on Environmental Quality, Executive Office of the President, 1989.

Cook, Thomas and Donald T. Campbell. Quasi-Experimentation; Design & Analysis Issues for Field Settings. Chicago: Rand McNally, 1979.

Cornfield, J. And J. W. Tukey. “Average Values of Mean Squares in Factorials,” Annals of Mathematical Statistics, 27: 907-949 (1956).

Covello, Vincent T., and others. “Communicating Risk Information to the Public,” Risk Communication: Proceedings of the National Conference on Risk Communication. Washington D.C.: The Conservation Foundation, 1987.

Covello, Vincent T., Sandman, Peter M., and Slovic, Paul. Risk Communication, Risk Statistics, and Risk Comparisons: A Manual for Plant Managers. Washington, D.C.: Chemical Manufacturers Association, 1988.

Covello, Vincent T. “Communicating Right-To-Know Information on Chemical Risks,” Environmental Science and Technology, 23: 1444-1449 (December 1989).

Creswell, John W. Research Design: Qualitative and Quantitative Approaches. Thousand Oaks, CA: Sage Publications, 1994.

DeSanctis, Gerardine. "Computer Graphics as Decision Aids: Directions for Research," Decision Sciences, 15: 463-487 (1984).

Edington, E. S. "Statistical inference and Nonrandom Samples," Psychological Bulletin, 66: 485-487 (1966).

Emory, William C. And Donald R. Cooper. Business Research Methods. 4th ed. Homewoood, IL: Richard D. Irwin, 1991.

Encyclopaedia Britannica. vol 5 Micropaedia. 15th ed. Chicago: Encyclopaedia Britannica, 1994.

Enrick, Norbert L. Handbook of Effective Graphic and Tabular Comunication. Hunington, NY: R. E. Krieger Publishing, 1980.

Environmental Protection Agency. A Citizen's Guide to Radon. 2nd ed. Document 402-K92-001. Washington: Government Printing Office, 1992.

Keppel, Geoffrey. Design and Analysis: A Researcher's Handbook. 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.

Lundgren, Regina E. Risk Communication: A Handbook for Communicating Environmental, Safety, and Health Risks. Columbus: Battelle Press, 1994.

MaKay and Villarreal. "Performance Differences in the Use of Graphic and Tabular Displays of Multivariate Data," Decision Sciences, 18-4: 535-545 (1987).

National Research Council. Improving Risk Communication. Washington D.C.: National Academy Press, 1989.

Neter, John, and William Wasserman and Michael H. Kutner. Applied Linear Statistical Models. 3rd ed. Boston: Richard D. Irwin, 1990.

Nunnally, J. C. Psychometric Theory. New York: McGraw-Hill, 1978.

Parsons, Robert. Statistical Analysis: A Decision-Making Approach. New York: Harper & Row, 1974.

Raymond, Neil V., Colonel, U.S. Air Force. Handout, Environmental Seminar. School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, 21 February 1995.

Riley, M. W. Sociological Research: A Case Approach. New York: Harcourt Brace Jovanovich, 1963.

Sandman, Peter M., and others. "High Risk or Low: How Location on a "Risk Ladder" Affects Perceived Risk," Risk Analysis, 14: 35-45 (January 1994).

Slovic, Paul, B. Fischhoff, and L Lichtenstein. "Characterizing Perceived Risk," in R. Kates, C. Hohenemser, and R. Kasperson (Eds.), Perilous Progress: Managing the Hazards of Technology. Boulder, CO: Westview.

Slovic, Paul. "Perception of Risk," Science, 236: 280-285 (1987).

Spector, Paul E. Summated Rating Scale Construction, An Introduction. Series/Number 07-082. Newbury Park: Sage Publications, 1992.

Szoka, Paller, Alan, and Kathryn, and Nan Nelson. Choosing the Right Chart: A Comprehensive Guide for Computer Users. San Diego: Integrated Software Systems Corporation, 1991.

Tan, Joseph K. H. And Izak Benbasat. "The Effectiveness of Graphical Presentation for Information Extraction: A Cumulative Experimental Approach," Decision Sciences, 24: 167-191 (1993).

Thal, Alfred E., Major, U.S. Air Force. Handout, Environmental Seminar. School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH, Novermber 1994.

Spyridakis, Jan H. "Conducting Research in Technical Communication: The Application of True Experimental Designs," Technical Communication, 39-4: 607-624 (November 1992).

Weinsetin, Neil D., Peter M. Sandman, and Nancy E Roberts. Communicating Effectively About Risk Magnitudes. Document EPA-230-08-89-064. Washington D.C.: U.S. Environmental Protection Agency, August 1989.

Weinsetin, Neil D., Peter M. Sandman, and Nancy E Roberts. Communicating Effectively About Risk Magnitudes, Phase Two. Document EPA-230-09-91-003. Washington D.C.: U.S. Environmental Protection Agency, August 1991.

Appendix A: Test Instrument

Thank you for participating in this study. The purpose of this research is to examine presentation formats and perceived risk. You will be asked to read a short brochure about a health hazard. In the brochure you will be given an *imaginary* test result. This *imaginary* test result, along with an accompanying chart, will be used to answer a few questions. The entire task will take approximately 10 minutes. You will not be paid for your participation.

As a participant in this experiment you have certain rights. The purpose of this document is to make you aware of your rights and to obtain your informed consent to certify your willingness to participate in this research study.

1) You have the right to stop participating in this experiment at any time. If you decide to do so, you should notify the experimenter immediately.

2) You have the right to withdraw your data from the experiment. Data are processed after all experimental runs are completed. All data are treated confidentially; therefore, if you wish to withdraw your data, you must do so immediately.

3) You have the right to be informed of the overall results of the experiment. If you wish to receive information about the results, a summary of the overall results of the experiment will be made available to you upon request free-of-charge. You may request a summary of results by including your address below your signature on the informed consent form and results will be sent to you after all data have been collected and analyzed.

If you have any questions please feel free to contact Captain Lonny Baker or Dr. Kim Campbell at the Air Force Institute of Technology, Wright-Patterson AFB, OH, or call (513) 255-2998.

CONSENT STATEMENT

I have read the above information and understand that participation is voluntary, refusal to participate will involve no penalty or loss of benefits to which I am otherwise entitled and I may discontinue participation at any time without penalty or loss of benefits to which I am entitled. My signature below means that I have freely agreed to participate in this research study.

____

Address (only if you wish to receive the overall results of the research):

Following this page is a three-page brochure . In the brochure you will be given an imaginary test result. Please read all the information in the brochure, then use your imaginary test result with the corresponding chart to answer the questions. If you have any questions before, during or after the test please notify the experimenter.

WHAT IS FIBRONITE ?

Fibronite is a mineral fiber found in rocks. There are several kinds of fibronite fibers, all of which are fire-resistant and not easily destroyed by natural processes. Because of its desirable qualities, fibronite has been used in a wide variety of products including appliances, ceilings, wall and pipe coverings, floor tiles, and some roofing materials.

WHY THE CONCERN ABOUT FIBRONITE ?

Although fibronite has many benefits for humans, it is also a very dangerous mineral. Breathing airborne fibronite fibers has been shown to cause: **(1)** Mintosis - a serious lung disease which can lead to disability and death; **(2)** Lung cancer - a disease that is incurable and almost always fatal; and **(3)** Mesothelioma - cancer of the lining of the lungs or abdominal cavities. The greater the exposure to fibronite, the more likely it is that one of these serious diseases will develop. Workers who handle or come into contact with fibronite on a daily basis are open to the greatest health risks.

HOW DOES FIBRONITE AFFECT US?

The danger arises when the fibronite fibers are released from the product or material. These fibers are so small they cannot be seen. They can float in the air for a long time and can pass through the filters of normal vacuum cleaners and get back into the air. Once inhaled, fibronite fibers can become lodged in tissue for a long time. After many years cancer or mintosis can develop. Cigarette smoking combined with fibronite exposure is especially hazardous.

Fibronite found in "friable" materials is most dangerous. Friable materials are materials that can be crumbled, pulverized, or reduced to powder by hand pressure. Materials containing fibronite which are sprayed on ceilings and walls are examples of friable materials. In contrast, vinyl fibronite floor tile is not usually friable. The fibronite fibers are firmly bound or sealed into the tile and can be released into the air only if the tile is cut, ground, or sanded.

WHERE IS FIBRONITE LIKELY TO BE FOUND IN THE HOME?

There are several areas in the home where fibronite problems are most likely to arise. These include:

- Wall construction material and pipe insulation, especially those dating between 1932 and 1968.
- Friable ceilings in buildings built or remodeled between 1940 and 1970.
- Material found in stoves and furnaces such as insulation and door gaskets.

Other fibronite-containing products that you may find in the home include:

- Patching compounds and textured paints (applied prior to 1967)
- Vinyl floor tiles and flooring.
- Roofing, shingles, and siding.
- Appliances with fibronite-containing parts or components, such as toasters, broilers, slow cookers, dishwashers, refrigerators, ovens, ranges, and clothes dryers.

Having significant amounts of fibronite in a residence is not rare. Many older residences in your area could create health problems for residents because of materials that may release fibronite fibers into the air.

HOW CAN I TELL IF I HAVE FIBRONITE IN MY RESIDENCE?

People who have worked frequently with fibronite (such as plumbers, and building or heating contractors) can often tell you whether or not material contains fibronite by looking at it.

If you suspect that you have a problem, you may also want to have an air sample taken to measure the number of fibronite fibers circulating inside your residence. To collect the sample, a laboratory will send a technician to your home. A pump is used to draw air from the room into a filter that will trap the fibronite. An electron microscope is used to count the sample. The test costs between $50 and $100, depending upon the laboratory technique used. The results of the test can be reported in units of "fibers per liter of air." This unit tells how much fibronite there is in one liter of air.

WHAT SHOULD I DO IF I HAVE A FIBRONITE PROBLEM?

If you discover that you have a fibronite problem, the best thing to do is to contact a contractor who has experience in the proper procedures for repairing and removing fibronite. There are special guidelines for handling fibronite materials. It is highly recommended that you hire an experienced contractor or get professional advice if you are thinking of doing the work yourself. Using improper techniques can make an existing problem much worse by contaminating the entire residence. For more information about identifying, testing, handling, and fixing fibronite problems call the Ohio Department of Health at (513) 285-6250.

Pretend that you have just had your residence tested for fibronite. The testing company tells you that you have a reading of **22 fibers per liter**. Try to answer the questions just as you would if you had actually received these results. FEEL FREE TO REFER TO THE BROCHURE WHEN ANSWERING THE FEEDBACK QUESTIONS.

INTERPRETING YOUR TEST RESULT

Use the chart on the next page of this brochure to interpret the imaginary test result you have been given. First, look down the left-hand column headed "Fibronite Level," and find the number nearest to your fibronite test result.

Next, move toward the right to the middle column headed "Extra Cancer Deaths." The number in this column tells you how many people are expected to die of cancer because of fibronite out of every 1000 people who live in a residence with the same fibronite level as yours.

Finally, move over to the right-hand column entitled "Comparison to Smoking Risk." It tells you how many cigarettes a day a person would have to smoke to have the same cancer risk as living with your fibronite level.

THE

RISK LADDER

APPEARS

ON

THIS

PAGE

FEEDBACK QUESTIONNAIRE

Please mark one and only one answer for each question. If none of the answers match your exact response, mark one that best matches how you feel. There are no right or wrong answers.

1. How would you describe the danger from your imaginary fibronite level

[ ] no danger
[ ] very slight danger
[ ] slight danger
[ ] moderate danger
[ ] serious danger
[ ] very serious danger

2. How likely do you think it is that continued exposure to your imaginary fibronite level would eventually have harmful effects? (Even though you may feel uncertain, please choose an answer to tell us what impression you got from the information you read.)

[ ] no chance
[ ] very unlikely
[ ] unlikely
[ ] moderate chance
[ ] likely
[ ] very likely
[ ] certain to happen

How do you think you would feel if your own residence actually had the fibronite level found by the imaginary test?

3. [ ] not at all concerned
[ ] slightly concerned
[ ] concerned
[ ] very concerned
[ ] extremely concerned

4. [ ] not at all frightened
[ ] slightly frightened
[ ] frightened
[ ] very frightened
[ ] extremely frightened

Once you have completed this page, please do not change any of your answers.

FOR CLASSIFICATION PURPOSES ONLY, please tell us:

a. Your sex: [ ] male [ ] female

b. Your age: __________

c. If applicable, what is your military rank or equivalent? __________________________

d. What is your career field? ____

e. How many years of experience do have in your career field? ______________________

f. Do you smoke? [ ] Yes [ ] No

g. How much schooling have you completed?

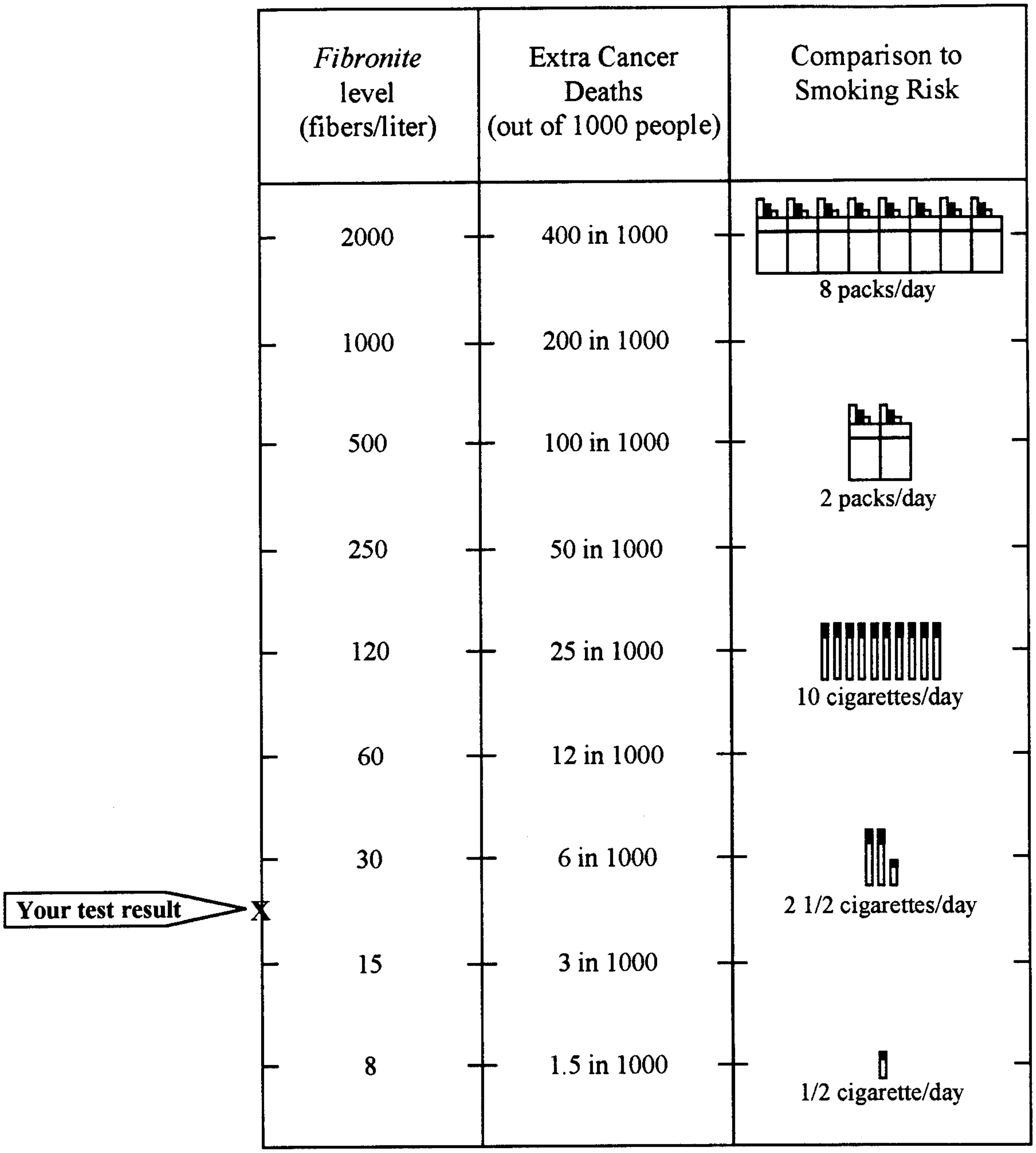
[ ] finished elementary school
[ ] some high school
[ ] finished high school
[ ] some college
[ ] finished 2-year college
[ ] finished 4-year college
[ ] some graduate study
[ ] graduate degree

h. Did you recognize the similarity of fibronite to asbestos? [ ] Yes [ ] No

Thank you, again, for your participation in this study. We must inform you that **fibronite is an imaginary hazard created solely for the purposes of this study**. Please do not become concerned about the presence of fibronite in your home. However, most of the information about fibronite which you read was taken from a brochure on asbestos. If you would like more information about the hazards associated with asbestos, you can contact the Department of Health in your state, the Environmental Protection Agency, or the National Institute for Environmental Health & Safety at 1-800-NIEHS-94.

Appendix B: Risk Ladders

INTERPRETING YOUR TEST RESULT:



Fibronite level (fibers/liter)	Extra Cancer Deaths (out of 1000 people)	Comparison to Smoking Risk
2000	400 in 1000	8 packs/day
1000	200 in 1000	
500	100 in 1000	2 packs/day
250	50 in 1000	
120	25 in 1000	10 cigarettes/day
60	12 in 1000	
30	6 in 1000	2 1/2 cigarettes/day
15	3 in 1000	
8	1.5 in 1000	1/2 cigarette/day

Figure 11. Location of Assigned Risk *Low*, No Action Standard

INTERPRETING YOUR TEST RESULT:

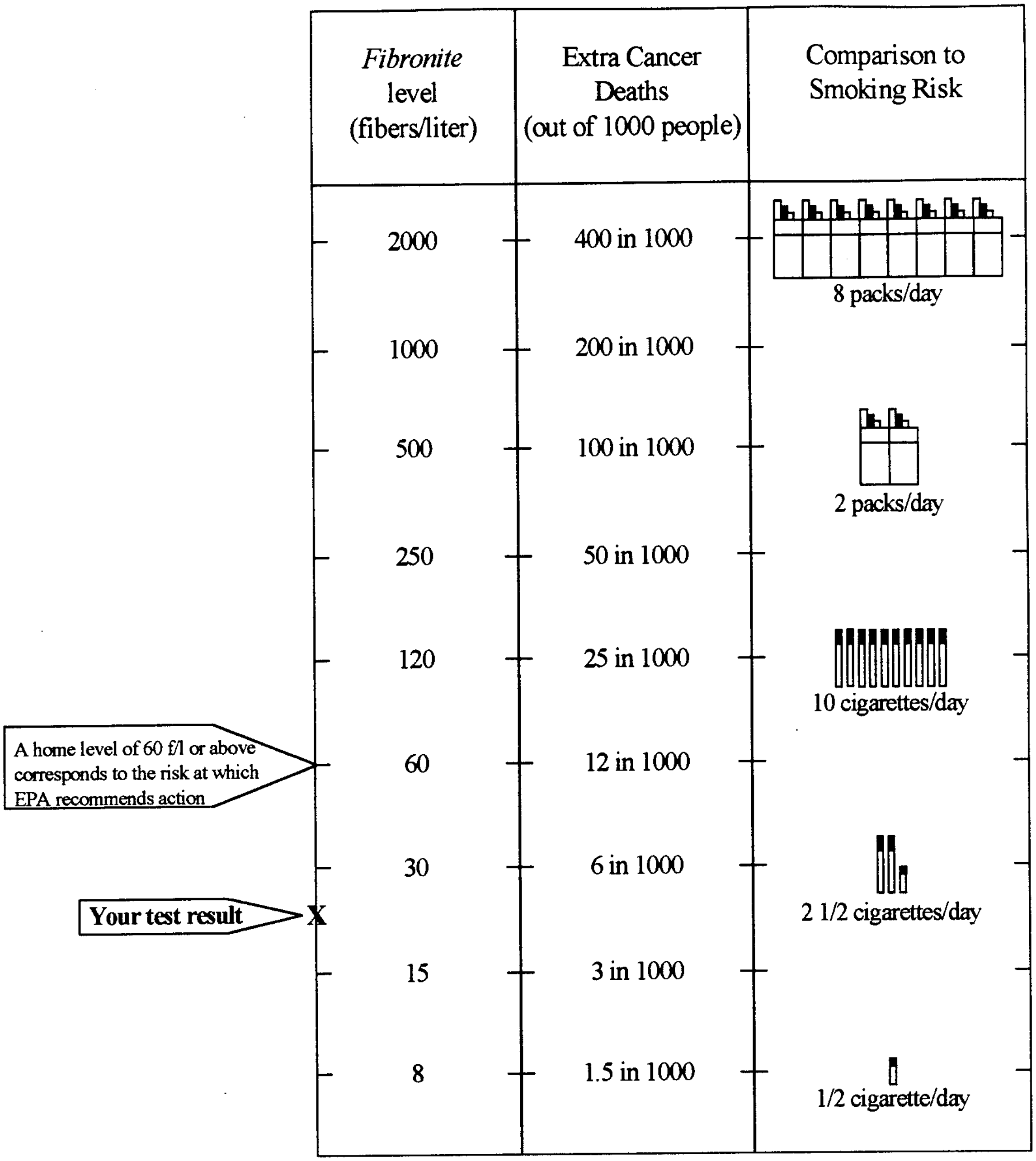


Figure 12. Assigned Risk *Low*, Action Standard *Above* Assigned Risk

INTERPRETING YOUR TEST RESULT:

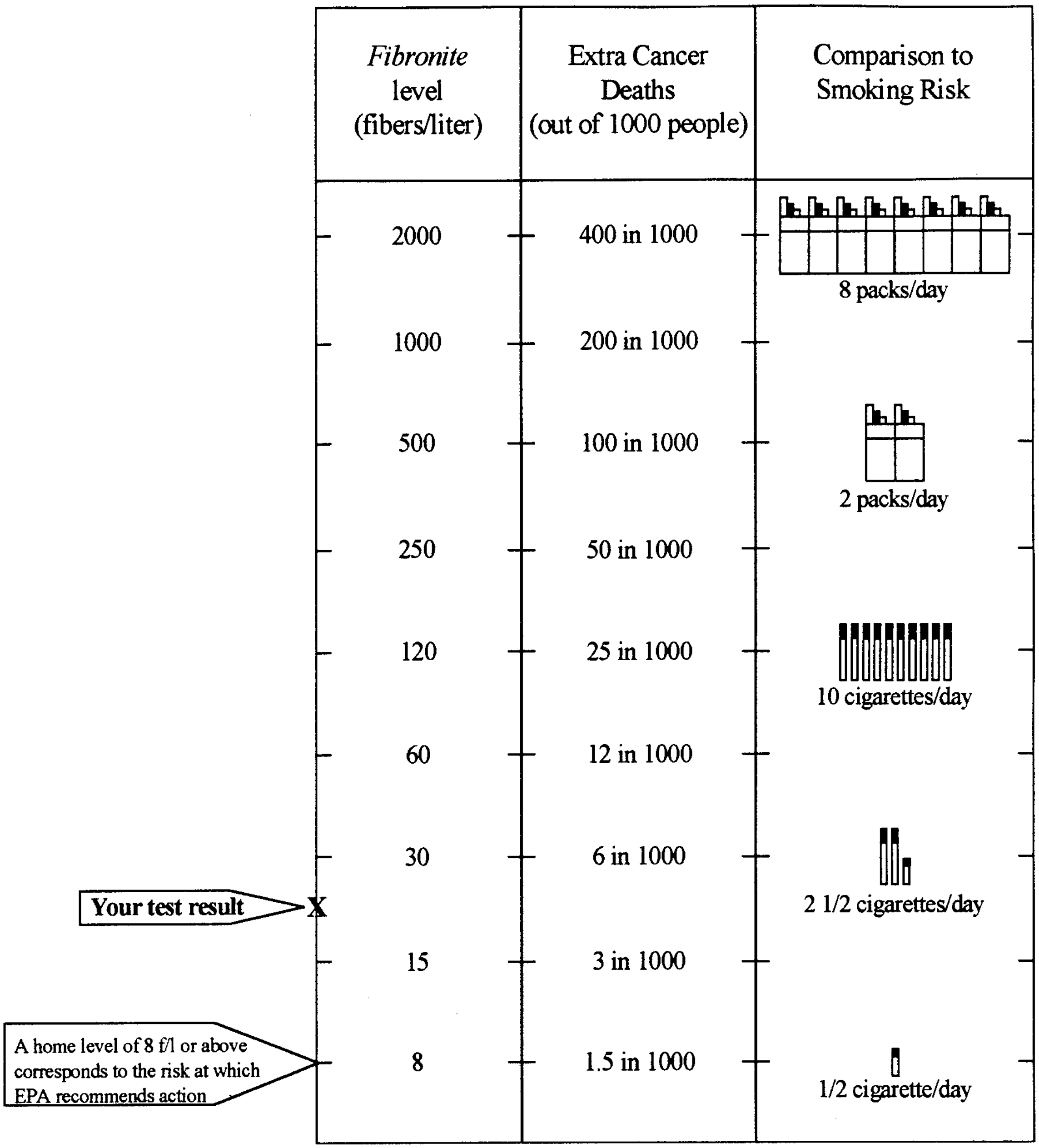


Figure 13. Location of Assigned Risk *Low*, Action Standard *Below* Assigned Risk

INTERPRETING YOUR TEST RESULT:

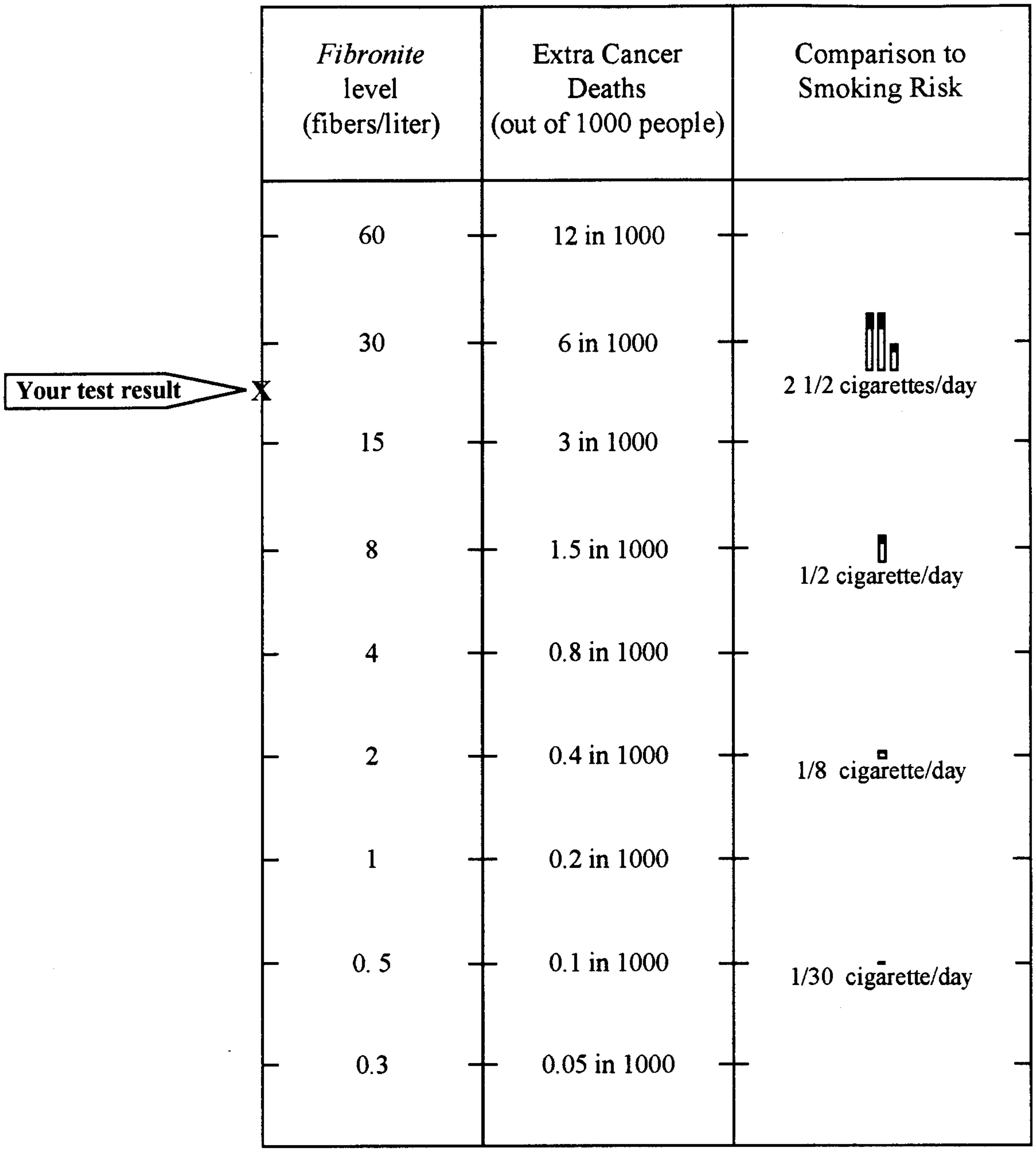
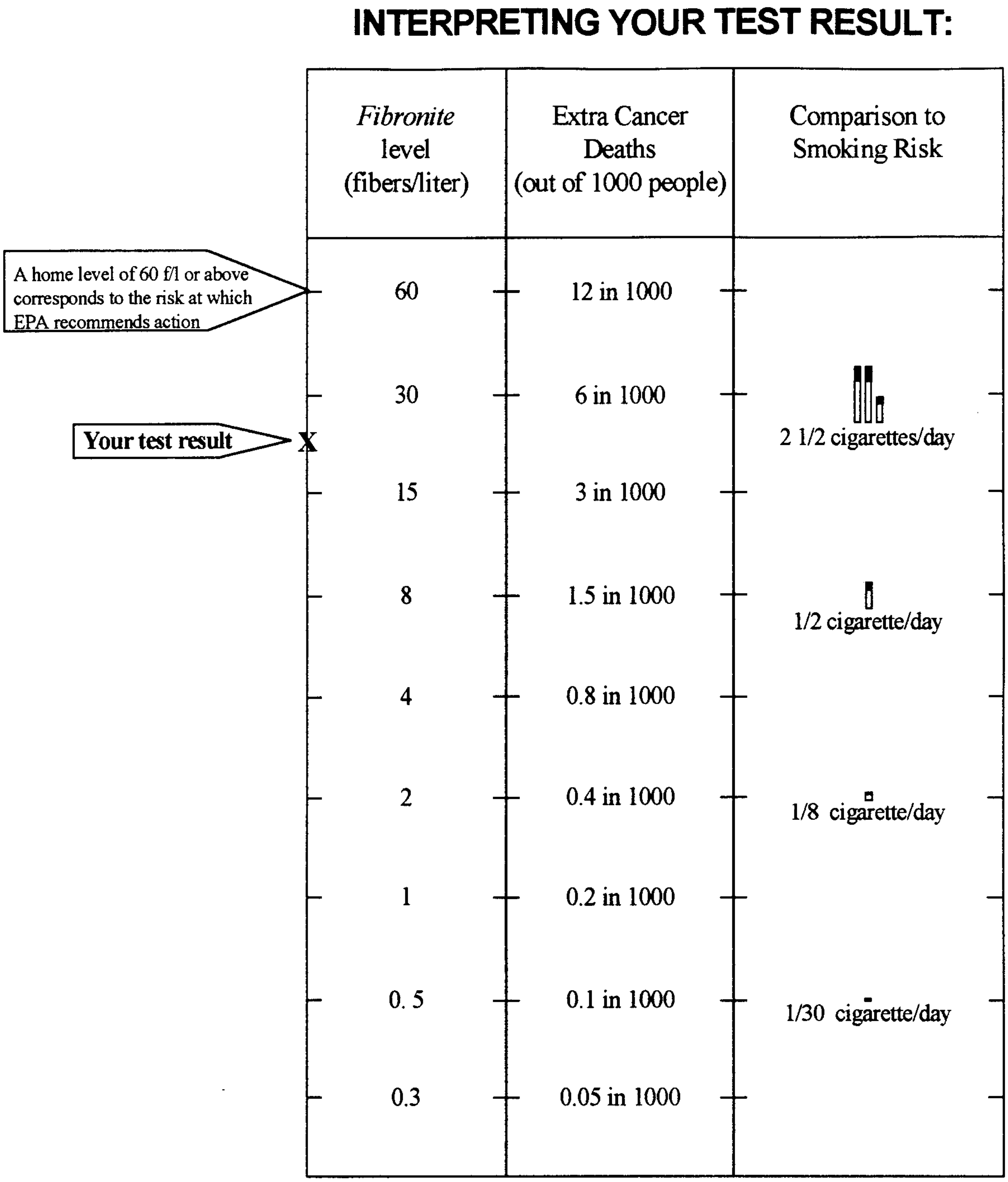


Figure 14. Location of Assigned Risk *High*, No Action Standard



INTERPRETING YOUR TEST RESULT:

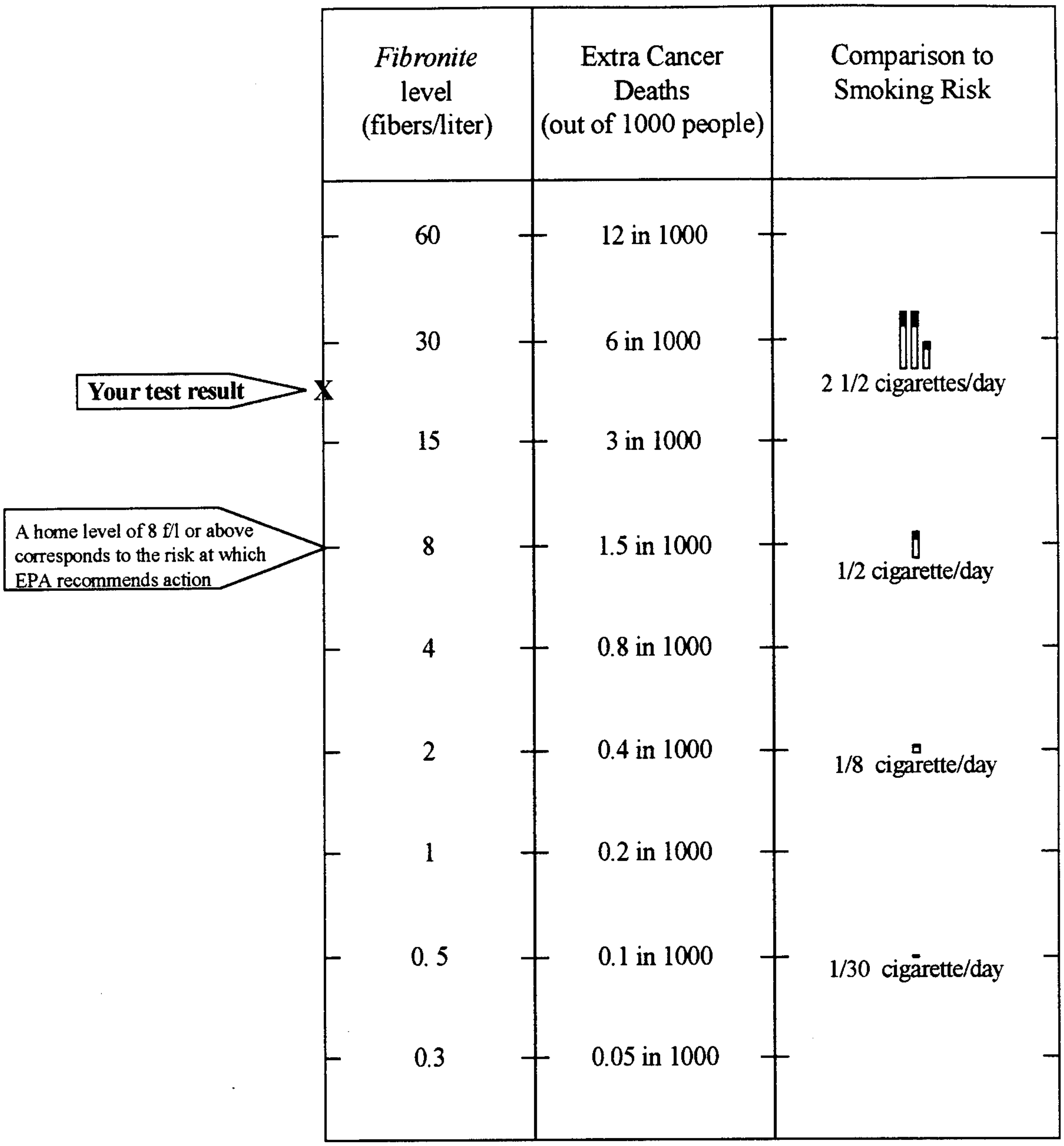
Fibronite level (fibers/liter)	Extra Cancer Deaths (out of 1000 people)	Comparison to Smoking Risk
60	12 in 1000	
30	6 in 1000	2 1/2 cigarettes/day
15	3 in 1000	
8	1.5 in 1000	1/2 cigarette/day
4	0.8 in 1000	
2	0.4 in 1000	1/8 cigarette/day
1	0.2 in 1000	
0. 5	0.1 in 1000	1/30 cigarette/day
0.3	0.05 in 1000	

A home level of 60 f/l or above corresponds to the risk at which EPA recommends action

Your test result X

Figure 15. Location of Assigned Risk *High*, Action Standard *Above* Assigned Risk

INTERPRETING YOUR TEST RESULT:



Fibronite level (fibers/liter)	Extra Cancer Deaths (out of 1000 people)	Comparison to Smoking Risk
60	12 in 1000	
30	6 in 1000	2 1/2 cigarettes/day
15	3 in 1000	
8	1.5 in 1000	1/2 cigarette/day
4	0.8 in 1000	
2	0.4 in 1000	1/8 cigarette/day
1	0.2 in 1000	
0. 5	0.1 in 1000	1/30 cigarette/day
0.3	0.05 in 1000	

Figure 16. Location of Assigned Risk *High*, Action Standard *Below* Assigned Risk

Appendix C: Characteristics of Sound Measurement

During the past few decades, the widespread acknowledgment of the importance of good measurement has led to the development of systematic and general approaches to measurement in the social sciences. As applied to the social sciences:

> . . . measurement is most usefully viewed as the process of linking abstract concepts to empirical indicants, as a process involving an explicit, organized plan for classifying the particular sense data at hand--the indicants--in terms of the general concept in the researcher's mind. (Riley, 1963:23)

Measurement, then, is a process which involves theoretical as well as empirical considerations.

> Empirically, the focus is on the *observable response*. Theoretically, interest lies in the *underlying unobservable* (and directly unmeasurable) *concept* that is represented by the response. Measurement focuses on the crucial relationship between the empirically grounded indicator(s)---that is, the observable response---and the underlying unobservable concept(s). When this relationship is a strong one, analysis of empirical indicators can lead to useful inferences about the relationships among the underlying concepts. In this manner, social scientists can evaluate the empirical applicability of theoretical propositions. (Carmines and Zeller, 1979:10-11)

Acknowledging the importance of sound measurement, how does one determine the extent to which a particular empirical indicator represents a given theoretical concept? In other words, what are the desirable qualities of any measuring procedure or instrument?

At the most general level, there are two basic properties of empirical measurements. They are reliability and validity (Cook and Campbell, 1979; Carmines and Zeller, 1979). "Fundamentally, *reliability* concerns the extent to which an experiment, test, or any measuring procedure yields the same results on repeated trials" (Carmines and

Zeller, 1979:11). *Valid(ity)* refers to the extent that (a test) measures what it purports to measure (Carmines and Zeller, 1979:12).

There are two major types of validity, internal and external, which need to be considered. Internal validity is concerned with the results of an experiment, and if there is a causal relationship between the variables. External validity is concerned with the observed causal relationship and if this relationship is generalizable to and across types of persons, settings, and times (Cook and Campbell, 1979:50, 71). While no experiment is ever completely free from threats to validity, it is important to try and reduce them. Equally important, researchers must be aware of and, acknowledge, the threats that may exist.

Internal Validity

Internal validity is also the ability of an experimental instrument to measure what it is suppose to measure. Some of the most common threats to internal validity follow:

(1) Subject History - All events that occur in the experimental environment except for the experimental conditions, should occur for all subjects.

(2) Subject Selection - To ensure that the selection and assignment of subjects to conditions does no result in biased groups, experimenters should identify one subject pool that contains subjects with similar backgrounds and then select all experimental subjects from that pool.

(3) Subject Maturation - If an experiment is conducted over time, there is a danger that subjects may have differing experiences outside the confines of the experiment and mature or change between experimental sessions.

(4) Testing. When an experiment involves multiple tests (such as pre-test, test, post-test), there is a danger that subjects will experience a learning effect which influences the results of the experiment.

(5) Instrumentation - Threats from instrumentation occur whenever experimental

instruments vary across conditions. Examples are software or hardware malfunctions. This threat can also occur when multiple test administrators, observers, or raters administer conditions or evaluate test performance.

(6) Statistical Regression - To avoid the threat of scores regressing to the mean, researchers should avoid selecting subjects on the basis of extreme scores or pre-measures.

(7) Experimental Mortality - This a problem that occurs when the composition of the original research group change during the course of the experiment.

(8) Expectancy - Subjects' or experimenters' expectations about experimental conditions can create results in and of themselves. This threat can easily occur when subjects are given too much information in the introduction to an experiment. (Cook and Campbell, 1979:51-55; Spyridakis, 1992:612-613)

These threats are usually taken care of by randomization (Cook and Campbell, 1979:56). This experiment used a completely randomized full factorial design, so most of the threats are dealt with sufficiently.

Subject history was not a concern because all subjects had relatively the same experience except for the experimental conditions. The only exception is that the three groups of subjects tested in three different classrooms. However, each room was approximately the same temperature, had the same lighting, and had the same seating arrangement. Each subject received the same written instructions, and the same test administrator. All printed materials were run from a single printer at the same time.

Subject selection and maturation were not a threat problems because "When respondents are randomly assigned to treatment conditions, each condition is similarly constituted on the average (no selection, maturation, or selection-maturation problems)" (Cook and Campbell, 1979:56).

Testing and instrumentation were not problems because each subject experienced the same testing conditions and experimental instrument. Also, there was only a single test which was administered by the same person.

Statistical regression was not a factor because no deliberate selection was made of high and low scorers on any tests or premeasures.

Due to the length of the test (10-15 minutes), the composition of the group did not change at any time. Therefore, experimental mortality was not a threat to validity.

Expectancy was not a threat because each subject received no instructions other than those written on the informed consent form. It was felt that the instructions did not give too much information about the study to bias the results.

External Validity

External validity refers to "... whether the findings will be representative and whether the results can be generalized to similar circumstances and subjects" (Creswell, 1994:134). The three major threats to external validity are as follows:

(1) Interaction of Selection and Treatment - Can a cause-effect relationship be generalized beyond the groups used to establish the initial relationship? The population the subjects are selected from may not be the population to which the results are to be generalized.

(2) Interaction of Setting and Treatment - Can a causal relationship obtained in one setting be obtained in another? For instance, can the results obtained in a factory be obtained in a military camp or on a university campus, or can the results obtained in an experimental setting be extrapolated to an actual setting? The experimental setting may have a biasing effect on the subjects' responses.

(3) Interaction of History and Treatment - To which periods in the past and future can a particular causal relationship be generalized? One may not get the same results on a special day, such as a holiday or a birthday, or on days with historical significance (such as the verdict of the O. J. Simpson trial), as one

would on a more mundane day. (Cook and Campbell, 1979:73-74):

"Problems of internal validity are amenable to solution by the careful design of experiments, but this is less true for external validity" (Emory and Cooper, 1991:428). External validity is ". . . the degree to which we can generalize findings beyond the present conditions of testing" (Keppel, 1991:484).

The interaction of selection and treatment may pose a threat to external validity. It is questionable as to whether the cause-effect relationships found in this experiment can be generalized beyond the group which was tested to establish the cause-effect relationship.

> One feasible way of reducing this bias is to make cooperation in the experiment as convenient as possible. For example, volunteers in a television-radio audience experiment who have come downtown to participate are much more likely to be atypical than are volunteers in an experiment carried door-to-door. (Cook and Campbell, 1979:73)

The 1991 EPA study was able to reduce this threat by sending the experiment out by mail and waiting for responses. While, in this experiment, participation was entirely voluntary, the subjects were more or less a 'captive' audience. However, based on the prior discussion about random and convenience samples, the results may be generalized to the population of the Air Force (at least for the military and civilian rank equivalents).

The interaction of treatment and setting may also pose a threat to external validity. Subjects who are tested in the classroom under the eyes of an experimenter may not behave similarly if they are allowed to complete the test at their own leisure in the convenience of their own home. "The solution . . . is to vary settings and to analyze for a causal relationship within each" (Cook and Campbell, 1979:74). Due to time, fiscal, and manpower constraints, this experiment necessitated a classroom testing environment.

It is doubtful that the interaction of history and treatment pose a significant threat to external validity. The three days over which the experiment was conducted had no important historical significance, either past or present. ". . . [S]olutions for short-term historical effects lie either in replicating the experiment at different times or in conducting a literature review to see if prior evidence exists which does not refute the causal relationship" (Cook and Campbell, 1991:74). Replication of the experiment was not feasible and no literature could be found which refute the causal relationships found in this study.

Reliability

Random error is involved in any type of measurement, no matter how refined the measuring techniques and instruments.

> . . . [*R*]*eliability* concerns the extent to which an experiment, test, or any measuring procedure yields the same results on repeated trials. The measurement of any phenomenon always contains a certain amount of chance error. The goal of error-free measurement--while laudable--is never attained in any area of scientific investigation." (Carmines and Zeller, 1979:11).

There are four basic methods for estimating the reliability of empirical measurements:

> (1) Retest Method - one of the easiest ways to estimate the reliability of empirical measurements give the same test to the same people after a period of time. One then obtains the correlation between scores on the two administrations of the same test.
>
> (2) Alternative-Form Method - similar to the retest method in that it also requires two testing situations with the same people. However, the same test isn't given on the second testing but an alternative form of the same test is administered.

(3) Split-Halves Method - can be conducted on one occasion. The total set of items is divided into halves and the scores on the halves are correlated to obtain an estimate of reliability. The halves can be considered approximations to alternative forms.

(4) Internal Consistency Method - requires a single test administration and provide an estimate of reliability for the given test administration. The most popular reliability estimate is given by Cronbach's alpha. (Carmines and Zeller, 1979:37-51):

Due to the time constraint for this thesis process and the availability of subjects, a duplication of this effort was practically impossible. The internal consistency method, using Cronbach's alpha, was used to assess reliability because it was easy to use and also because it is one of the most popular and accepted methods for assessing reliability. Additionally, Cronbach's alpha provides a conservative estimate of a measure's reliability (Carmines and Zeller, 1979:45).

Appendix D: Calculations for Homogeneity of Variance

The Hartley test is based on the H distribution where the distribution of H depends on the number of populations r and the common number of degrees of freedom df (Neter et al., 1990:619). If each of the r sample variances s_i^2 has the same number of degrees of freedom df, the statistical hypotheses can be stated as:

$$H_o\text{:}\ \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2 = \sigma_6^2$$

$$H_a\text{:}\ \text{not all } \sigma_i^2 \text{ are equal}$$

The test statistic is:

$$H = \frac{\max(s_i^2)}{\min(s_i^2)}$$

The appropriate decision rule for controlling the risk of making a Type I error at α is:

$$\text{If } H \leq H(1-\alpha\,;\, r,\, df)\text{, conclude } H_o$$

$$\text{If } H > H(1-\alpha\,;\, r,\, df)\text{, conclude } H_a$$

where $H(1-\alpha\,;\, r,\, \text{df})$ is the $(1-\alpha)100$ percentile of the distribution of H when H_o holds, for r populations and df degrees of freedom for each sample variance (Neter et al., 1990:619).

For this study $r = 6$ and $df = \text{n-1} = 13\text{-}1 = 12$. The level of significance was controlled at $\alpha = .05$. The value of H_{critical} to which $H_{observed}$ was compared, was found using the *Percentiles of H Distribution* table in the appendix of Neter, Wasserman, and Kutner (Neter et al., 1990:1154).

Using the table, it was found that $H(0.95;\ 6,\ 12) = 5.72$, and

$$H = \frac{\max(s_i^2)}{\min(s_i^2)} = \frac{14.638}{7.102} = 2.061$$

Therefore, since $H_{observed} \leq H_{critical}$, it was concluded that the variances of each treatment group were equal.

Appendix E: Raw Data

The columns labled Q1, Q2, Q3, and Q4 contain the replies to questions 1, 2, 3, and 4, respectively, from the test questionnaire. The columns labled Qa, Qb, Qc, Qd, Qe, Qf, Qg, and Qh contain the replies to questions a-h, respectively, in the feedback questionnaire. Column Qd contains the career field of each subject. Each career field is entered as a one or two letter designation. The meaning of each designation is as follows:

Table 6
Career Field Designations for Raw Data Tables

Designation	Career Field
A	Acquisition
C	Cartography
E	Engineer
F	Finance/Auditor
L	Logistics
M	Managemetn
Mu	Munitions
Mx	Maintenance
Pm	Program Manager
Q	Quality Assurance
S	Supply
T	Supply

Table 7
Raw Data for Treatment One

LOCATION-Low, FORMAT-No Action Standard

SUBJECT	Q1	Q2	Q3	Q4	Qa	Qb	Qc	Qd	Qe	Qf	Qg	Qh
1	5	7	5	5	M	29	N	M	8	N	8	Y
2	4	4	5	3	F	40	N	L	7	N	5	Y
3	4	6	4	3	M	28	O3	S	5	N	6	Y
4	4	5	3	2	M	24	O2	Mx	2	N	7	Y
5	4	5	3	2	M	23	O1	Mx	1	N	6	N
6	5	5	5	3	M	31	O3	L	9	N	7	Y
7	2	3	2	2	M	52	N	T	25	N	6	Y
8	4	5	3	2	F	38	N	F	15	N	7	Y
9	3	5	3	2	M	32	O3	Mx	10	N	8	Y
10	5	6	4	4	M	35	E6	A	17	Y	4	Y
11	4	5	3	2	M	29	O1	M	0.5	N	7	Y
12	3	4	3	2	M	26	O2	M	4	N	7	Y
13	5	6	3	3	M	38	E7	Mx	17	N	5	N

Table 8
Raw Data for Treatment Two

LOCATION-Low, FORMAT-Action Standard Above

SUBJECT	Q1	Q2	Q3	Q4	Qa	Qb	Qc	Qd	Qe	Qf	Qg	Qh
14	3	3	2	1	M	37	O4	M	12	N	8	Y
15	3	3	2	1	M	48	G11	C	10	N	8	Y
16	5	6	5	4	F	37	G12	L	8	N	5	Y
17	2	3	2	1	F	37	G12	M	4	N	3	Y
18	1	2	3	2	M	26	O1	A	1	N	7	Y
19	5	6	4	3	M	40	O3	L	10	N	8	Y
20	2	4	2	2	M	23	O1	L	1	N	6	Y
21	3	3	2	2	M	43	E7	L	25	N	5	Y
22	2	2	4	3	F	35	G9	Q	7	N	4	Y
23	4	5	3	3	F	40	G9	M	2	N	8	Y
24	4	5	3	2	M	42	O5	E	18	N	8	Y
25	3	3	2	2	F	40	N	L	11	N	7	Y
26	3	4	3	2	M	55	N	M	30	N	7	Y

Table 9
Raw Data for Treatment Three

LOCATION-Low, FORMAT-Action Standard Below												
SUBJECT	**Q1**	**Q2**	**Q3**	**Q4**	**Qa**	**Qb**	**Qc**	**Qd**	**Qe**	**Qf**	**Qg**	**Qh**
27	4	5	5	2	M	44	O4	A	12	N	8	Y
28	2	3	2	2	M	40	O5	M	16	N	8	Y
29	5	3	2	2	M	41	G12	L	5	N	4	Y
30	1	1	3	1	F	41	G11	M	6	N	4	N
31	4	5	4	1	M	36	N	E	13	N	6	N
32	3	4	2	1	M	42	O5	A	18	Y	8	Y
33	3	4	4	3	M	47	N	F	23	N	8	Y
34	3	4	3	2	M	24	O2	E	2	N	7	Y
35	3	3	3	1	M	24	O2	A	2	N	7	Y
36	3	4	3	3	M	35	E7	T	18	Y	3	Y
37	4	4	3	2	M	35	O4	A	8	N	8	Y
38	3	4	3	1	M	48	O4	L	26	N	5	Y
39	2	4	2	1	M	38	G12	E	12	N	8	Y

Table 10
Raw Data for Treatment Four

LOCATION-High, FORMAT-No Action Standard												
SUBJECT	Q1	Q2	Q3	Q4	Qa	Qb	Qc	Qd	Qe	Qf	Qg	Qh
40	3	4	3	2	M	43	O5	Mx	8	N	8	Y
41	5	5	4	4	M	42	O4	S	11	N	8	Y
42	4	7	5	4	M	41	N	L	20	N	5	Y
43	5	6	4	2	F	47	G1	L	18	N	5	Y
44	4	6	5	3	F	44	G1	L	10	N	6	Y
45	3	4	3	2	M	38	O3	Mx	16	N	8	Y
46	5	5	4	3	M	34	E6	T	16	N	3	Y
47	5	5	3	2	M	50	N	L	10	N	3	Y
48	4	4	4	3	M	48	G1	L	4	Y	7	Y
49	4	3	2	1	M	38	O3	L	19	N	8	Y
50	4	6	4	4	M	32	O7	T	14	N	4	Y
51	5	6	5	5	M	33	O2	Mx	8	N	8	Y
52	6	7	5	1	M	32	E7	A	10	Y	5	Y

Table 11
Raw Data for Treatment Five

LOCATION-High, FORMAT-Action Standard Above

SUBJECT	Q1	Q2	Q3	Q4	Qa	Qb	Qc	Qd	Qe	Qf	Qg	Qh
53	2	3	2	1	M	26	G9	E	3	N	7	Y
54	4	4	4	2	M	43	O4	L	12	N	7	Y
55	4	5	3	3	F	39	G1	L	14	N	8	Y
56	4	3	3	2	M	37	E7	T	10	N	5	Y
57	4	4	3	2	M	43	E8	A	25	N	4	Y
58	3	4	2	1	M	41	E7	T	18	N	5	Y
59	3	3	2	1	M	42	N	L	20	Y	4	Y
60	6	6	5	4	F	36	E6	T	17	Y	5	Y
61	3	6	4	4	M	34	E6	A	14	N	5	Y
62	4	4	3	2	F	47	N	L	18	Y	5	Y
63	4	6	4	2	M	31	O3	Mx	7	N	7	Y
64	4	4	3	2	M	40	O3	Mx	3	N	8	Y
65	5	5	4	3	M	31	O3	E	4	N	8	Y

Table 12
Raw Data for Treatment Six

LOCATION-High, FORMAT-Action Standard Below

SUBJECT	Q1	Q2	Q3	Q4	Qa	Qb	Qc	Qd	Qe	Qf	Qg	Qh
66	5	5	5	3	M	24	G1	E	2	N	7	Y
67	4	4	2	1	M	25	O2	E	2	N	7	Y
68	2	5	3	2	M	35	E7	Mu	18	N	5	Y
69	4	4	3	2	M	27	O3	Mx	7	N	8	Y
70	5	6	4	3	F	33	N	F	10	N	6	Y
71	5	6	4	4	F	31	E6	S	11	N	5	Y
72	4	5	5	4	M	29	E3	Mx	6	Y	4	Y
73	5	5	5	4	F	44	G1	A	10	N	4	Y
74	5	6	5	5	M	38	E7	Mx	17	Y	5	Y
75	4	6	4	2	M	26	O2	A	3	N	6	Y
76	4	5	4	4	M	37	E7	Mx	18	N	4	N
77	5	5	4	3	M	47	O4	Mx	18	Y	8	Y
78	3	3	3	1	M	37	G1	E	13	N	7	Y

Appendix F: SAS Code and Output

SAS Code and Output for Cronbach's Alpha and Intercorrelation

SAS Code

```
options linesize=80;
data palmer;
infile data;
input Q1 1 Q2 2 Q3 3 Q4 4;
proc corr alpha;
var Q1 Q2 Q3 Q4;
proc freq;
tables Q1 Q2 Q3 Q4;
```

SAS Output

```
                        The SAS System
               10:28 Tuesday, September 12,1995

                        Correlation Analysis

          4 'VAR' Variables:  Q1        Q2        Q3        Q4


                         Simple Statistics

Variable          N        Mean     Std Dev         Sum   Minimum  Maximum

Q1               78     3.76923     1.09216   294.00000         1        6
Q2               78     4.52564     1.25589   353.00000         1        7
Q3               78     3.41026     1.01208   266.00000         2        5
Q4               78     2.41026     1.08635   188.00000         1        5


                      Correlation Analysis

                   Cronbach Coefficient Alpha
               for RAW variables         : 0.876504
               for STANDARDIZED variables: 0.878600


                       Raw Variables                 Std. Variables

 Deleted        Correlation                   Correlation
 Variable        with Total         Alpha      with Total          Alpha

 Q1               0.708489       0.851613       0.701686       0.858266
 Q2               0.777339       0.826808       0.776577       0.828993
 Q3               0.774684       0.829259       0.776383       0.829070
 Q4               0.692551       0.857610       0.697155       0.860000

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / N = 78
```

	Q1	Q2	Q3	Q4
Q1	1.00000 0.0	**0.72396** 0.0001	**0.61548** 0.0001	**0.51867** 0.0001
Q2	0.72396 0.0001	1.00000 0.0	**0.67619** 0.0001	**0.61092** 0.0001
Q3	0.61548 0.0001	0.67619 0.0001	1.00000 0.0	**0.71902** 0.0001
Q4	0.51867 0.0001	0.61092 0.0001	0.71902 0.0001	1.00000 0.0

Q1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	2	2.6	2	2.6
2	8	10.3	10	12.8
3	19	24.4	29	37.2
4	28	35.9	57	73.1
5	19	24.4	76	97.4
6	2	2.6	78	100.0

Q2	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1	1.3	1	1.3
2	2	2.6	3	3.8
3	14	17.9	17	21.8
4	21	26.9	38	48.7
5	21	26.9	59	75.6
6	16	20.5	75	96.2
7	3	3.8	78	100.0

Q3	Frequency	Percent	Cumulative Frequency	Cumulative Percent
2	16	20.5	16	20.5
3	28	35.9	44	56.4
4	20	25.6	64	82.1
5	14	17.9	78	100.0

Q4	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	16	20.5	16	20.5
2	31	39.7	47	60.3
3	17	21.8	64	82.1
4	11	14.1	75	96.2
5	3	3.8	78	100.0

SAS Code and Output for the ANOVA

SAS Code

```
option pagesize=60;
option linesize=80;
option nocenter;
OPTIONS NOOVP CC=CR;
title 'Capt Baker - Thesis Analysis';
filename NEW 'PAUL2.dat';
data PAUL2;
infile NEW;
input SUBJ   1-3  LOCATE$ 4-8  FORMAT$ 9-14  CI 15-16;
proc sort DATA=PAUL2;
by SUBJ LOCATE FORMAT;
proc means DATA=PAUL2 noprint;
by SUBJ LOCATE FORMAT;
var CI;
output OUT=PAUL2MN MEAN=MEANCI;
proc univariate normal plot;
var MEANCI;
proc anova DATA=PAUL2MN;
classes SUBJ LOCATE FORMAT;
model MEANCI=SUBJ(LOCATE FORMAT) LOCATE|FORMAT;
test h=LOCATE e=SUBJ(LOCATE FORMAT);
test h=FORMAT e=SUBJ(LOCATE FORMAT);
test h=LOCATE*FORMAT e=SUBJ(LOCATE FORMAT);
means LOCATE;
means FORMAT;
means LOCATE/tukey e=SUBJ(LOCATE FORMAT);
means FORMAT/tukey e=SUBJ(LOCATE FORMAT);
means LOCATE*FORMAT;
```

SAS Output

```
Capt Baker - Thesis Analysis           10:36 Thursday, September 14, 1995
Univariate Procedure
Variable=MEANCI

                              Moments

                    N                78  Sum Wgts          78
                    Mean       14.11538  Sum             1101
                    Std Dev    3.810378  Variance    14.51898
                    Skewness    0.06647  Kurtosis    -0.91613
                    USS           16659  CSS         1117.962
                    CV          26.9945  Std Mean     0.43144
                    T:Mean=0   32.71689  Pr>|T|        0.0001
                    Num ^= 0         78  Num > 0           78
                    M(Sign)          39  Pr>=|M|       0.0001
                    Sgn Rank     1540.5  Pr>=|S|       0.0001
                    W:Normal   0.957603  Pr<W          0.0361

                           Quantiles(Def=5)

                     100% Max         22        99%        22
                      75% Q3          17        95%        21
                      50% Med         14        90%        19
                      25% Q1          11        10%         9
                       0% Min          6         5%         8
                                                 1%         6

                      Range           16
                      Q3-Q1            6
                      Mode            10

                              Extremes

                             Lowest Obs     Highest Obs
                                 6(30)         20(42)
                                 8(53)         21(51)
                                 8(18)         21(60)
                                 8(17)         21(74)
                                 9(59)         22(1)


Variable=MEANCI
                    Stem Leaf                     #  Boxplot
                      22 0                         1      |
                      21 000                       3      |
                      20 00                        2      |
                      19 0000                      4      |
                      18 00000000                  8      |
                      17 00000000                  8   +-----+
                      16 0000                      4   |     |
                      15 0000                      4   |     |
                      14 00000000                  8   *--+--*
                      13 0000000                   7   |     |
                      12 00000000                  8   |     |
                      11 000                       3   +-----+
                      10 00000000                  8      |
                       9 000000                    6      |
                       8 000                       3      |
                       7                                  |
                       6 0                         1      |
```

Normal Probability Plot

```
22.5+                                              +++*
    |                                          **+*
    |                                        **++
    |                                     ***+
    |                                 ****+
    |                             ****+++
    |                            **+++
    |                           **+
14.5+                        ***+
    |                      ***
    |                   ****
    |                 +**
    |              ****
    |         *****+
    |     * **+++
    |       ++
 6.5+ *  +++
     +----+----+----+----+----+----+----+----+----+----+
         -2        -1         0        +1        +2
```

Analysis of Variance Procedure
Class Level Information

```
Class      Levels    Values

SUBJ           78    1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
                     23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41
                     42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
                     61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78

LOCATE          2    high low

FORMAT          3    above below none


Number of observations in data set = 78

Dependent Variable: MEANCI
                                       Sum of            Mean
Source                    DF          Squares          Square   F Value     Pr > F

Model                     77     1117.9615385      14.5189810         .          .

Error                      0                .               .

Corrected Total           77     1117.9615385

        R-Square              C.V.         Root MSE           MEANCI Mean

        1.000000                 0                0             14.115385


Source                    DF         Anova SS     Mean Square   F Value     Pr > F

SUBJ(LOCATE*FORMAT)       72     826.15384615     11.47435897         .          .
LOCATE                     1     120.62820513    120.62820513         .          .
FORMAT                     2     124.69230769     62.34615385         .          .
LOCATE*FORMAT              2      46.48717949     23.24358974         .          .

Tests of Hypotheses using the Anova MS for SUBJ(LOCATE*FORMAT) as an error
term
```

```
Source                      DF          Anova SS      Mean Square    F Value   Pr > F

LOCATE                       1      120.62820513     120.62820513      10.51   0.0018

Tests of Hypotheses using the Anova MS for SUBJ(LOCATE*FORMAT) as an error
term

Source                      DF          Anova SS      Mean Square    F Value   Pr > F

FORMAT                       2      124.69230769      62.34615385       5.43   0.0063

Tests of Hypotheses using the Anova MS for SUBJ(LOCATE*FORMAT) as an error
term

Source                      DF          Anova SS      Mean Square    F Value   Pr > F

LOCATE*FORMAT                2       46.48717949      23.24358974       2.03   0.1394


Level of          ------------MEANCI-----------
LOCATE       N       Mean                SD

high        39    15.3589744        3.57246177
low         39    12.8717949        3.67193830


Level of          ------------MEANCI-----------
FORMAT       N       Mean                SD

above       26    12.7692308        3.75561119
below       26    13.7692308        3.82944061
none        26    15.8076923        3.31082142

Tukey's Studentized Range (HSD) Test for variable: MEANCI

NOTE: This test controls the type I experimentwise error rate, but generally
      has a higher type II error rate than REGWQ.

Alpha= 0.05  df= 72  MSE= 11.47436
Critical Value of Studentized Range= 2.819
Minimum Significant Difference= 1.5292

Means with the same letter are not significantly different.

     Tukey Grouping              Mean       N  LOCATE

                  A           15.3590      39  high

                  B           12.8718      39  low

Tukey's Studentized Range (HSD) Test for variable: MEANCI

NOTE: This test controls the type I experimentwise error rate, but generally
      has a higher type II error rate than REGWQ.

Alpha= 0.05  df= 72  MSE= 11.47436
Critical Value of Studentized Range= 3.384
Minimum Significant Difference= 2.2483

Means with the same letter are not significantly different.
```

```
        Tukey Grouping              Mean      N  FORMAT

                    A            15.8077     26  none
                    A
               B    A            13.7692     26  below
               B
               B                 12.7692     26  above


Level of    Level of              -------------MEANCI------------
LOCATE      FORMAT         N         Mean                  SD

high        above         13      13.6923077           3.59130003
high        below         13      16.0769231           3.45112393
high        none          13      16.3076923           3.32627458
low         above         13      11.8461538           3.82635932
low         below         13      11.4615385           2.66506362
low         none          13      15.3076923           3.35123399
```

Vita

Capt Lonny P. Baker was born on March 8, 1963 in Tucson, Arizona. He graduated from Greenbrier East High School in Fairlea, West Virginia in 1981. He entered the United States Air Force Academy and graduated with a Bachelor of Science degree in International Relations, Russian Area Studies in 1985. His first assignment was at Francis E. Warren AFB as a missile launch officer. His next assignment was at Langley AFB where he served in the command post of the 1st Fighter Wing. His next assignment was at Bitburg AB, Germany where he was the Air Base Operability flight commander. When Bitburg AB closed in 1994, he entered the School of Engineering, Air Force Institute of Technology to pursue a Masters degree in Engineering and Environmental Management.

Permanent Address:
P.O. Box 119
Fort Spring, WV 24936

REPORT DOCUMENTATION PAGE

[illegible]

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
December 1995

3. REPORT TYPE AND DATES COVERED
Master's Thesis

4. TITLE AND SUBTITLE
How Location and an Action Standard on a Risk Ladder Affect Perceived Risk

5. FUNDING NUMBERS

6. AUTHOR(S)
Lonny P. Baker, Captain, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology, WPAFB OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GEE/ENV/95D-02

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Efforts to explain risk magnitudes often rely on a risk ladder in which exposure levels and associated risk estimates are arranged with low levels at the bottom of the ladder and high levels at the top. An experiment was conducted to test the hypotheses that perceived threat varies with the location of a subject's assigned level on the risk ladder, and that perceived threat varies with the presence of an action standard on the risk ladder. Air Force Institute of Technology Professional Continuing Education students were asked to assume a particular level of a hypothetical hazard in their residences, to read a brochure explaining the risks, and then to complete a questionnaire. The study found that subjects with an assigned risk level three-quarters of the way up the ladder perceived higher levels of risk than subjects with an assigned risk level one-quarter of the way up the ladder. The study also found that the presence of an action standard on a ladder may significantly affect perceived risk. Subjects who received a risk ladder without an action standard perceived a higher level of risk than subjects who received a risk ladder with an action standard placed above assigned risk level. These findings suggest that risk response can be shaped by effectively presented data alone.

14. SUBJECT TERMS
environment, communication, risk communication, risk ladders, risk magnitudes, action standards, graphics

15. NUMBER OF PAGES
102

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
Unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT
Unclassified

20. LIMITATION OF ABSTRACT
UL