AFIT/GSM/LAS/97S-3

THE EFFECTS OF EXTRANEOUS PRESENTATION GRAPHICS ON DECISION MAKING

THESIS

Thomas S. Tingley, Captain, USAF

AFIT/GSM/LAS/97S-3

AFIT/GSM/LAS/97S-3

THE EFFECTS OF EXTRANEOUS PRESENTATION GRAPHICS ON DECISION MAKING

THESIS

Presented to the Faculty of the Graduate School of

Logistics and Acquisition Management of the

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Systems Management

Thomas S. Tingley, B.S.

Capt, USAF

September 1997

Approved for public release, distribution unlimited

ACKNOWLEDGMENTS

Many people provided outstanding support of this research effort. Particularly, I would like to thank the thesis committee for their guidance and direction, without which this thesis would not have been possible. Likewise, the PCE instructors who enthusiastically cooperated in the data collection process, helped to provide one of the most essential ingredients in any experiment - test subjects. Much appreciation is extended to all the PCE students, AFIT graduate students, and AFIT instructors who voluntarily participated in this study. Those who participated in the pilot study were especially helpful in developing and refining the test instrument. Finally, I would like to extend deepest gratitude to my wife for her exceptional patience, understanding, and support.

Thomas S. Tingley

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

List of Figures

Page

AFIT/GSM/LAS/97S-3

ABSTRACT

Almost every professional presentation includes some form of graphics. Many previous researchers have analyzed the relative effectiveness of graphical decision-aids, such as pie charts or bar graphs. Very little research, however, has been conducted to determine the effects of extraneous graphics, such as pictures and logos, on decision-making. This study employed an experiment to determine whether extraneous graphics hinder or enhance a decision-maker's ability to obtain the pertinent information from a professional presentation. Other factors such as gender and graphics training were also analyzed for any influence on the results obtained. The results of the study indicated a statistically significant effect on information extraction accuracy due to the use of extraneous graphics, but an effect was not discovered regarding decision accuracy, or decision confidence. Information extraction accuracy was found to be inversely related to the level of extraneous graphics employed in a presentation. Additionally, the effect on extraction accuracy was found to be more predominant for females and for people without prior experience or training (either formal or informal) in graphics applications. The tests also showed that most people prefer a "medium" amount of extraneous graphics rather than a large amount or no extraneous graphics.

THE EFFECTS OF EXTRANEOUS PRESENTATION GRAPHICS ON DECISION MAKING

I. INTRODUCTION

General Issues

The utilization of graphics in professional presentations has become routine over the last decade. All the popular word processing and presentation software packages include pre-made drawings and images (commonly referred to as "clipart") as well as the capability to produce original graphics. Many types of graphics such as pie charts, bar graphs, and X-Y plots have been shown to enhance the decision making process because these graphics can summarize large amounts of pertinent information in a small space. However, many professional presentations also include extraneous graphics, such as clipart images, which are primarily used to enhance the aesthetic value of the presentation. For the purpose of this study, extraneous graphics are defined as any non-essential images, drawings, pictures, logos, diagrams, or cartoons added to a presentation. These graphics are "non-essential" because they do not provide additional information, and not required for communicating the pertinent information of a presentation.

Some authors, such as Needleman, have claimed that the presentation format used to present the idea being advocated is critical to the decision-making process, especially for those presentations geared toward persuading the audience to adapt a particular point of view. "Regardless of whom you are making a presentation to, the purpose of a presentation isn't to make your point, but to sell an audience on your point of view!" (Needleman, 1993: 15). Other authors claim that extraneous graphics are "chart junk," which may distract the viewer from obtaining the information in the presentation (Tufte, 1983: 107). If the main points of a presentation are obscured by the packaging, the decision-maker may have been misled or distracted.

According to Peeler, successfully conveying the point of view, or message, of a presentation depends on two primary conditions. The first is the believability of the presentation, and the second is the accurate extraction of information by the decision-maker (Peeler, 1996: 2). To ensure successful transmission and reception of the message, "you not only have to understand what information you need to present, but also the best way to present it so that it reinforces your final goal, not obscures it" (Needleman, 1993: 15). The presenter, therefore, must choose the appropriate type and amount of graphics to capture the decision-maker's attention and facilitate accurate extraction of the information without distracting or adversely influencing the decision.

Specific Issue

Many forms of extraneous graphics are readily available in most popular presentation software packages, such as Microsoft PowerPoint®, as well as from a multitude of Internet sources. Many professionals in the Department of Defense (DoD) and in the private sector, seem to make extensive use of these extraneous graphics throughout their presentations. The problem is that most people that use extraneous graphics do not know what impact, if any, the graphics may have on decision-making.

Many previous researchers have analyzed the effectiveness of graphical decision-aids, such as pie charts or bar graphs. Very little research, however, has been conducted to determine the effects of extraneous graphics, such as clipart, pictures, and logos, on the decision-making process. The use of extraneous graphics may help to attract the viewer's attention, aid in the extraction and recall of important information, and possibly make a presentation more believable due to its "professional" appearance. On the other hand, extraneous graphics may simply distract the viewer and prevent the accurate extraction of important information.

The researchers studying the effectiveness of graphical decision-aids have reported conflicting results. Some researchers contend that many forms of graphics do indeed enhance the decision-making process (Horton, 1991: 12), while others consider some forms of graphics to be a waste of space (Tufte, 1983, 107). Still others maintain that empirical evidence neither supports nor rejects either position (Schaubroeck and Muralidhar 1991: 127, Jarvenpaa, 1986: 3). Most of the studies on this subject analyzed the effectiveness of tables versus graphs. The researchers based their conclusions on experiments which presented information to viewers in either graphic or tabular formats. The results from both subject groups were then compared to determine if any significant differences were observed. These experiments focused on which format (tabular or graphical) is more suited for data presentation. Although none of these experiments specifically addressed extraneous graphics, one researcher, Edward Tufte, offers his opinions regarding "chartjunk" (Tufte, 1983: 107).

Tufte characterizes chartjunk as, "the interior decoration of graphics" (107). By Tufte's definition, extraneous graphics are a form of chartjunk used to enhance the aesthetic value of a presentation, not the information it contains. He states that, "data graphics should draw the viewer's attention to the sense and substance of the data, not to something else," therefore, "a large share of ink on a graphic should present data-information" (91). Tufte proposes that the only "ink" (or surface area, for computer displays) used on a presentation slide should be employed to convey information. Any ink that does not provide additional information should be omitted from the presentation. Tufte's positions, however, are presented without the support of empirical evidence. It is quite possible that some forms of extraneous graphics can actually stimulate viewer interest, enhance information extraction accuracy, aid memory recall, and have a positive influence on decision-making.

One form of extraneous graphics are "background graphics," the effects of which were analyzed in a 1996 thesis effort by David Peeler. Background graphics are color schemes or pictures that fill in the background behind the essential information of a presentation. The pre-made background "Design Templates" available in Microsoft PowerPoint® are an example of background graphics. Peeler's research concluded that background graphics have no significant effect on decision accuracy or decision confidence (1996: 65). However, background graphics, by definition, are *behind* the information and are probably the least obtrusive form of extraneous graphics. This thesis effort concentrates on the more predominant pictures, logos, and icons displayed in the foreground of a presentation. These types of extraneous graphics are more likely to catch the viewer's attention, and therefore, more likely to influence decision-making. To avoid confusion with Peeler's work, the term "extraneous graphics" will not include background graphics for the purpose of this study.

As the capability to insert many forms of extraneous graphics into professional presentations has increased over the past several years, most presentations contain at least some form of extraneous graphics. Figure 1 shows a simple information slide without extraneous graphics, while Figures 2, 3, and 4 demonstrate various forms of extraneous graphics. Figure 2 shows the use of organizational logos, which are added to the same information that is in Figure 1. Figure 3 shows clipart added to the same informational slide, and Figure 4 illustrates the use of both organizational logos and clipart on the same slide.

Cockpit Upgrade Project

- **Voice Recognition System**
  - **Emergency situation response**
  - **Weapon system control**
- **Contract Status**
  - **$7 Million over budget**
  - **2 Months behind schedule**

Figure 1. Informational Slide without Extraneous Graphics

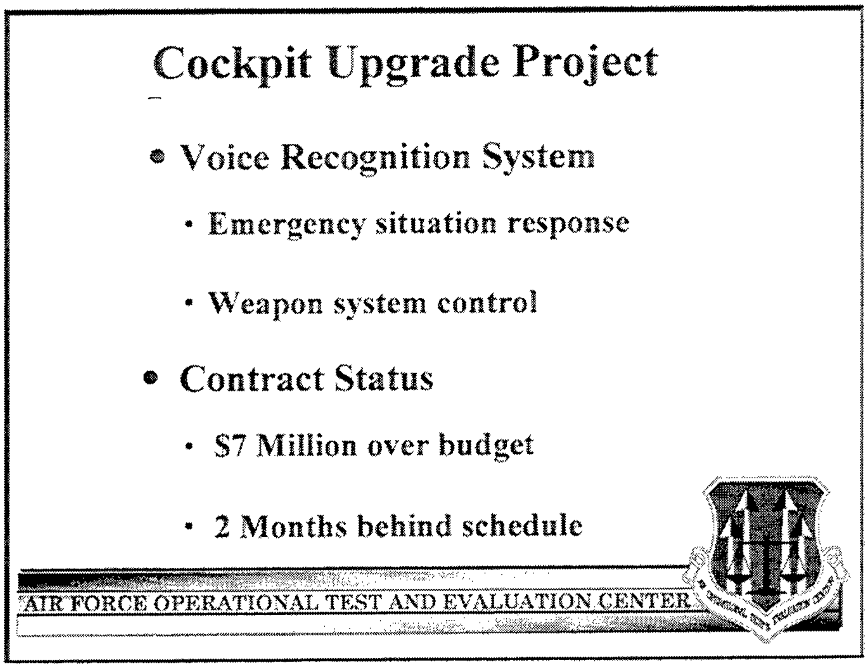


Figure 2. Informational Slide with Organizational Logo

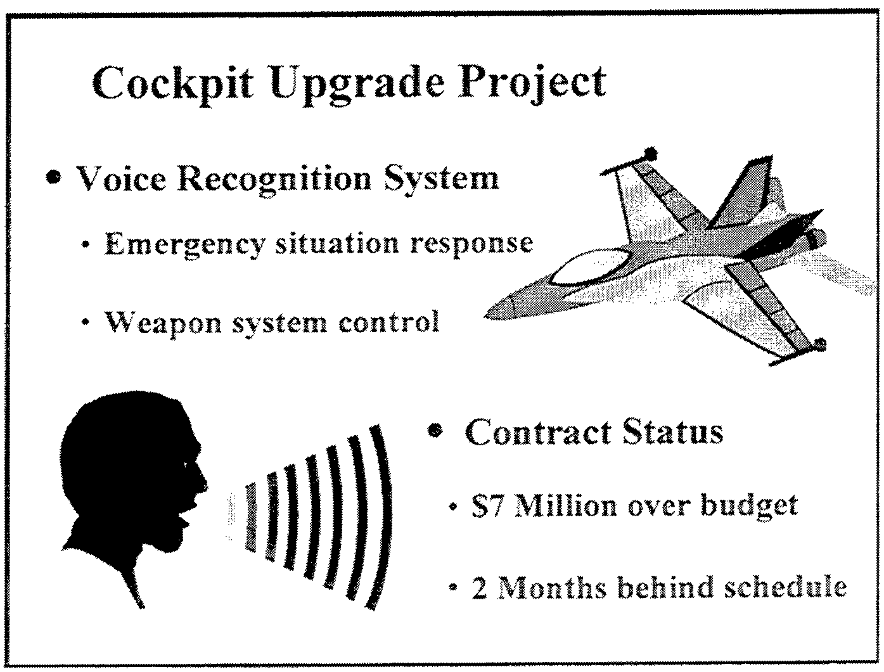


Figure 3. Informational Slide with Clipart

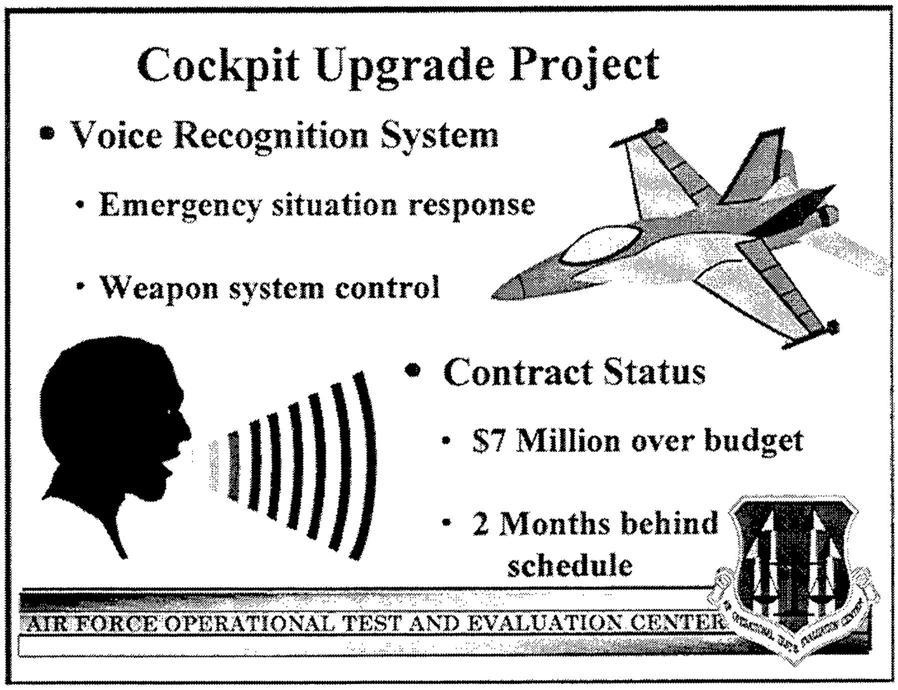


Figure 4. Informational Slide with Organizational Logos and Clipart

Comparing Figures 1 and 4, it is obvious that extraneous graphics can have dramatic effects on the appearance of an informational slide. Both slides contain the same original information, but Figure 4 certainly looks different than Figure 1. Which one has a more professional appearance? Does it appear that the presenter put more effort into developing the slide with more graphics, thus, making it more believable? Which one is the viewer more likely to remember? Which one facilitates accurate information extraction? Does either format enhance or hinder decision-making? Due to the individual nature of human perception and interpretation, a single answer to these questions for every possible viewer is unlikely. However, as an increasing number of presentation software users become proficient in the use of graphics options, some research should be conducted in order to answer these and other related questions concerning the influence of graphics applications on the DoD population.

Many important decisions, which are based on briefings or presentations, are made every day in the DoD. Many of these decisions have significant impact on the national defense. Within the hierarchical nature of the DoD decision-making structure, decisions made by lower level managers or commanders impact those at higher levels. Presentations used to aid in the decision-making process are used at each successive layer of the hierarchy. Many times the person receiving a briefing at one level, will become the briefer at the next level. If extraneous graphics influence decision performance in different ways for different individuals, then the information presented at subsequent hierarchical levels may have varying influential effects on different decision-makers. In other words, the clipart used to gain the attention of the colonel may irritate the general. If the information presented to DoD decision-makers at all levels was free of influential factors, these pitfalls could be avoided. However, the trend within the DoD, as well as in the civilian sector, appears to include more and more extraneous graphics. Therefore, a better understanding of the potential influential effects of these extraneous graphics is necessary.

The results obtained from graphics research could be used to improve the quality of decisions throughout the DoD. Research into the use of extraneous graphics will provide DoD personnel with information regarding the potential effects of such usage on decision-makers. Using the results of this research, briefers would be able to present information in formats that are less likely to adversely affect the proper decisions indicated by the data. Likewise, informed decision-makers would be able to understand the potential effects extraneous graphics may have on their decision-making ability, and allow them to make unbiased decisions which are based solely on the relevant information.

Previous research which has focused on the different effects of various graphical formats on decision-making, has not concentrated on the influencing effects of extraneous graphics. Therefore, the purpose of this thesis is to determine if extraneous graphics enhance or impede the accurate extraction of information, and to determine if extraneous graphics affect the decision-maker's ability to reach the appropriate decision indicated by the information presented.

Research Questions

To achieve the objectives of this study three research questions were identified:

Q1. Do extraneous graphics affect information extraction accuracy?

Q2. Do extraneous graphics affect decision accuracy?

Q3. Do extraneous graphics affect decision confidence?

Figure 5 illustrates the comparison of an informational slide without extraneous graphics, and the same informational slide with extraneous graphics. In this comparison, as in the experiment used to test these research questions, the original information is in the same position and format on both slides. The second slide (Panel B) simply has clipart pictures inserted around the informational graph.

Panel A Panel B

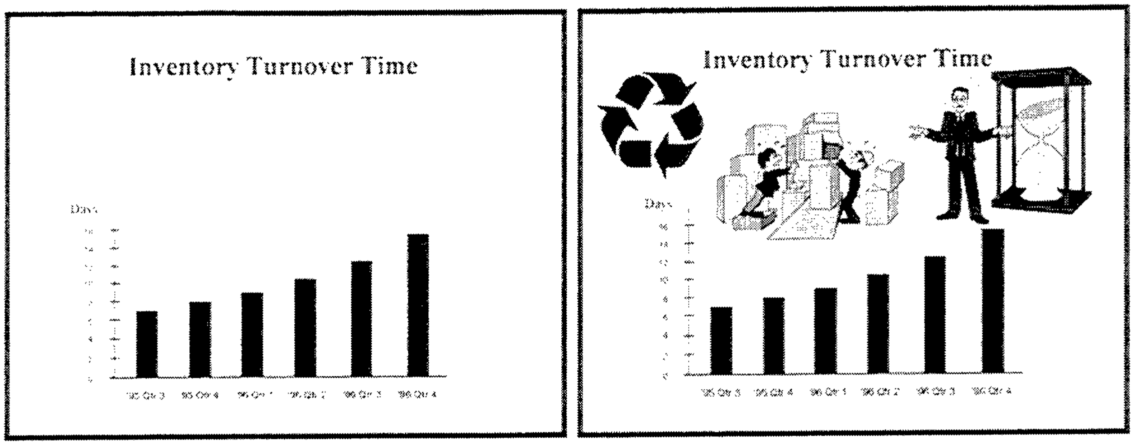


Figure 5. Comparison of Extraneous Graphics Application

The first research question will be employed to determine if the use of extraneous graphics affects the viewer's ability to accurately obtain the important information from a presentation. It is possible that limited use of extraneous graphics which are related to the subject information may enhance accurate information extraction, whereas extensive use of extraneous graphics are likely to distract the viewer and impede accurate information extraction. The null hypothesis would be that extraneous graphics have no effect on the viewer's ability to accurately obtain the pertinent information from a presentation. Therefore, there are three possible answers to the first research question: (1) extraneous graphics aid in information extraction; (2) extraneous graphics hinder information extraction; or (3) extraneous graphics have no effect on information extraction.

The second research question addresses whether or not the use of extraneous graphics influences the decision-maker's ability to reach the proper decision indicated by the information provided. As with the first research question, it is possible that some extraneous graphics may help the viewer to make the proper decision, but overuse of extraneous graphics will probably hinder the decision-making process. The null hypothesis would be that there is no effect on the decision-making process regardless of the presence or absence of extraneous graphics. Therefore, there are also three answers to the second research question: (1) extraneous graphics enhance decision accuracy; (2) extraneous graphics hinder decision accuracy; or (3) extraneous graphics have no effect on decision accuracy.

The third research question is used to determine whether the decision-maker's confidence in his or her decision can be influenced by the presence of extraneous graphics. This influence (if it occurs) is likely due to two causes. The first cause is related the first research question, since the decision-maker's decision is likely to be based on the information he or she obtains from the presentation. If decision-makers are uncertain of how well they were able to accurately extract the necessary information, their decision confidence is likely to be affected. The second cause is related to the "believability" of the presentation. Preliminary analysis has shown that viewers are more likely to believe (or feel confident with) the information in a presentation that is accompanied by extraneous graphics simply because it appears that the presenter has put more work into the presentation, and thus, knows what he or she is talking about. Again, there are three possible answers to this research question: (1) extraneous graphics increase decision confidence; (2) extraneous graphics decrease decision confidence; or (3) extraneous graphics have no effect on decision confidence.

In addition to these three research questions, several areas for sensitivity analysis are appropriate. Some previous research has found indications that gender and training (in graphical applications) have an influence on the results obtained. Previous research by MacKay and Villarreal indicates that gender may influence how well graphical information is interpreted (1987: 544). A background in graphics training (or the lack thereof) has also been identified as a possible influential factor on the comprehension of graphical presentations (DeSanctis, 1984: 477). Another potential influential factor for sensitivity analysis is colorblindness. It is possible that those who are colorblind may perceive and/or interpret the information differently than those who are not. These three factors are addressed as investigative questions for sensitivity analysis:

Q4: Does gender affect Q1, Q2, or Q3?

Q5: Does graphics training affect Q1, Q2, or Q3?

Q6: Does colorblindness affect Q1, Q2, or Q3?

Analysis will be done to determine if the influential effects (if any) found in the three research questions, differ based on the gender, level of graphics training, or colorblindness of the decision-makers.

Synopsis

This chapter begins with a discussion of the general issues associated with the use of extraneous graphics. The general issues were narrowed to identify the specific issues related to the application of extraneous graphics within presentations. Figures 1 through

5 illustrate the differences between presentation formats that omitted or included extraneous graphics. Following the specific issues to be addressed, the research questions employed to study these issues were presented and discussed.

Chapter II, the literature review, examines the research done to date regarding the use of graphics in conjunction with decision-making. Several studies which are related to the current research topic are analyzed and discussed. Chapter III, the methodology, explains the research hypothesis and the experimental design employed to test the hypotheses. Also included, are discussions regarding the validity of the design, its applicable concepts and constructs, as well as the construction and administration of the experimental item. Chapter IV, the analysis and findings, presents the results of the experiment. Chapter V, the conclusion, discusses the experimental results as they relate to the hypotheses presented in Chapter I. The results are also discussed in the context of their impact within the DoD. Finally, some suggestions are made regarding possible topics for future research concerning extraneous graphics.

II. Literature Review

This research is a follow-on effort to a 1996 AFIT thesis, by David Peeler. As a result, the focus of the literature review will be on Peeler's thesis followed by the related research. Because this is a follow-on effort using a similar experimental design, much of the literature review and methodology are similar to that of Peeler's thesis.

The views of graphics software manufacturers will be discussed, followed by an analysis of the previous research. A substantial number of previous research efforts have analyzed the use of graphical formats as replacements for tables or narrative text. Far fewer studies have been conducted concerning the use of graphs as decision-making tools. Very little research has been accomplished which specifically addressed the effects of extraneous graphics on decision-making.

Graphics Software

Computer graphics software programs have become a standard means of communication in today's business environment, as well as in the DoD. Software packages provide a plethora of extraneous graphics, and the software manufacturers boldly proclaim that colorful presentations employing many forms of graphics will have an enhanced aesthetic value and positively influence perceptions and decisions (Microsoft, 1995: 1, Claris, 1993: W-1).

Microsoft, Claris, and Apple create software for IBM compatibles and Macintosh systems. The messages these software producers are sending in their accompanying user's manuals, imply that graphics can be used to enhance the visual presentation of information, and thus, improve the capability of the decision-maker. Microsoft states that a "well-placed graphic can transform a plain-looking slide into a compelling visual message" (Microsoft, 1995: 272). Claris proposes that graphics produce presentations with a professional appearance, implying that presentations without graphics are unprofessional-looking (Claris 1993: W-1). Apple states that "you can create an even stronger visual framework by employing the two primary visual elements – graphics and type" (Apple, 1991: 26).

Regarding the relation between graphics and decision-making, the makers of Crystal Ball® software claim, "Crystal Ball is a user-friendly, graphically oriented forecasting and risk analysis program that takes the uncertainty out of decision-making." Furthermore, "With Crystal Ball, you will become a more confident, efficient, and accurate decision-maker" (Decisioneering, 1993: 1). Since the volume of available graphical software packages will only continue to grow, the actual effects of graphics on decision-making must be investigated. If graphics do indeed influence decision-making, decision-makers should become educated regarding the potential implications of these influential effects.

Previous Research

Numerous studies have been conducted to analyze the use of graphs versus tables regarding data comparison, interpretation, extraction speed, accuracy, performance, and preference. However, much less research has been accomplished concerning the possible effects of graphics on decision-making. Although extraction speed, accuracy, and performance are factors that affect decision-making, "they do not individually possess the all-inclusive essence of decision-making" (Peeler, 1986: 15). Of the reviewed studies, Peeler's was the only one that integrated these measures and tasks involved in decision-making into a decision problem. The experiment employed in Peeler's study required the use of presented data to achieve more than data extraction - the viewer was required to become a "decision-maker" by reaching a decision based on the information provided in the presentation. The other research efforts certainly help to establish the influence of graphics on answering a specific question regarding a particular slide, but do not attempt to use a sequence of graphics to inform a decision-maker, who must subsequently utilize the information obtained in order to reach a decision (Peeler, 1996: 17).

To summarize the results of pertinent studies regarding the use of graphics, a table is provided. Table 1 contains the names of the researchers responsible for their respective studies in the left-most column. The second and third columns reflect the independent and dependent variables for each study. The right column provides a short summary of the research results. In studies involving more than one dependent variable, the variables and corresponding results are numbered. Most of the information provided in Table 1 was obtained from a similar table in Peeler's thesis (1996: 16). The authors Schaubroeck and Muralidhar, as well as Peeler have been added to the table.

Table 1. Summary of Previous Graphics Research

Authors	Independent Variable	Dependent Variable	Results
Tullis, 1981	Narrative vs. tables vs. graphs	1. Attitude 2. Speed 3. Accuracy 4. Preference	1. Graphics better 2. Graphics better 3. No effect 4. Graphics better
Zmud, 1983	Tables vs. Graphs; Task complexity	Decision quality	Graphs better for low complexity tasks; tables best for high complexity
Corbone & Gorr, 1985	Graphs vs. Enhanced Graphics	Decision accuracy	No effect
Blocher, et al., 1986	Tables vs. Graphs; Task complexity	Decision quality	Graphs better for low complexity tasks; tables best for high complexity
Dickson, et al., 1986	Tables vs. Graphs; Bar graphs/Line graphs	1. Interpretation accuracy 2. Decision quality	1. No effect/No effect 2. No effect/Line graphs better
Jarvenpaa & Dickson, 1986	Tables vs. Graphs	1. Information retrieval 2. Information recall 3. Message comprehension 4. Recognizing trends 5. Recall large amounts of data	1. No effect 2. No effect 3. No effect 4. Graphs better 5. Graphs better
Davis, 1989	Tables vs. Graphs	Performance	Graphs better only when visual cues aid in answering questions
DeSanctis & Jarvenpaa, 1989	Tables vs. Graphs	1. Accurate interpretation 2. Incorporate into accurate judgments 3. Confidence	1. Tables better 2. Graphs better 3. No effect
Larkin, 1990	High integrity vs. Misleading graphs	Decision quality	Integrity graphs better
Schaubroeck & Muralidhar, 1991	Tables vs. Graphs	Decision accuracy	No effect
Barber & Dunn, 1992	Iconic vs. Traditional graphs	1. Accurate interpretation 2. Impression	1. No difference 2. Traditional better
Latin & Villanueva, 1994	3D graphs vs. 2D graphs vs. Tables	Decision accuracy	No effect
Peeler, 1996	Background graphics	1. Decision accuracy 2. Decision confidence	1. No effect 2. No effect

Peeler. In his 1996 AFIT thesis, Peeler studied the effects of graphics on decision-making. His research concentrated on a form of extraneous graphics referred to as "background graphics," which are defined as pictures or color schemes that fill in the background area behind the essential information presented to the decision-maker (Peeler, 1996: 1). An example of Peeler's application of background graphics is illustrated in Figure 6. The slide on the right contains the same graphic and information as the slide on the left. The only differences between the two slides are the addition of a background template and the accompanying color changes automatically implemented by the presentation software. Peeler's research is a first attempt to study the influential effects of extraneous graphics, and is a precursor to this thesis. The methodology and experimental design of this thesis are very similar to that of Peeler's.

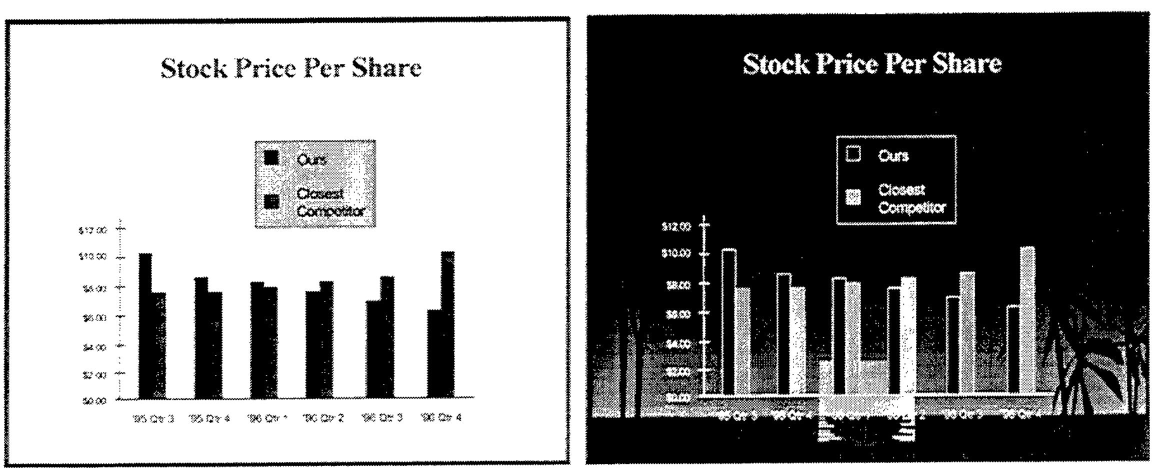


Figure 6. No Background Graphics vs. Background Graphics

Peeler did not find conclusive evidence that background graphics have a significant impact on decision-making accuracy or confidence. Background graphics, however, are probably the least obtrusive of the extraneous graphics since the important information is overlaid on top of (or in front of) the graphic. Other types of extraneous graphics such as pictures or images prominently displayed in the foreground may have more effect on the decision-making process. Regardless of his results, Peeler provides an important assessment, "the outcome of a presentation depends on two main ingredients: the believability of the presentation, and the accurate extraction of information by the decision-maker" (Peeler, 1996: 2).

Schaubroeck and Muralidhar. Schaubroeck and Muralidhar conducted a meta-analysis regarding the relative effects of tabular versus graphic display formats on decision-making. For their meta-analysis, the findings from 21 studies were quantitatively reviewed. Several of the studies analyzed by Schaubroeck and Muralidhar are individually reviewed in the following paragraphs. The results of their meta-analysis showed that sampling error and differential range restriction accounted for variability in the average differences between tabular and graphic presentations, and that there was no difference in the effects of the two formats (Schaubroeck and Muralidhar, 1991: 127). The authors also claimed that their results should discourage further investigation into the effects of tabular versus graphic formats "using the conventional methodology" (1991: 127). The studies that Schaubroeck and Muralidhar analyzed, however, did not seem to actually address decision-making (as discussed in the introductory paragraphs of this section). Due to the nature of the research, which was tables versus graphs, the study also did not address the effects of extraneous graphics.

Davis. Davis studied the effects of presentation modes (tables vs. graphs) and graph complexity on decision-making. He hypothesized that the more complex a table or graph, the less efficient the decision-maker will be. He also hypothesized that the more complex a table or graph, the less accurate will be the decisions reached by utilizing it. For the purpose of his study, Davis defined complexity as the amount of information that must be examined to reach a conclusion (Davis, 1989: 497).

Although his study of presentation modes showed inconclusive results regarding efficiency and effectiveness, further analysis seemed to show that "different forms of presentations are best for different tasks" (1989: 497). Therefore, Davis concluded that no one particular form of problem presentation, tabular or graphic, is best for all decision situations and the choice of presentation forms must take the decision task into consideration. Davis also proposes that his findings help explain the inclusive results of prior research efforts.

Latin and Villanueva. In their 1994 AFIT thesis, Latin and Villanueva studied the differing effects of three-dimensional graphs, two-dimensional graphs, and tables have on decision making. Specifically, they analyzed how these different presentation formats influence decision accuracy. The results of their research showed no significant differences in decision accuracy. Therefore, they concluded that the decision-maker is able to accurately extract the necessary information and reach the proper decision, regardless of the graphical format in which the information is presented (Latin and Villanueva, 1994: 47).

DeSanctis and Jarvenyaa. DeSanctis and Jarvenyaa studied the ability to make accurate decisions based on the format in which data is presented. Specifically, they tested whether "forecast accuracy will be better when data is displayed in a graphical format than when data is displayed in a numeric format" (DeSanctis and Jarvenyaa, 1989: 511). To do so, they developed a theoretical framework to combine extraction accuracy and judgments. The results of their study indicate that forecast accuracy was better using graphics while numeric tables provide more accurate interpretation in regards to identifying and communicating specific data values.

Corbone and Gorr. Corbone and Gorr supported the theory regarding the positive effects of graphics on judgment-type tasks, and thus hypothesized that improved graphics would further improve the decision-makers judgment capability. Improved or enhanced graphs refer to those that provide additional decision cues or redundant data to aid the decision-maker in reaching the appropriate conclusion. To test this hypothesis, they developed an experiment in which numerical references were added to bar charts to aid in interpretation accuracy, as shown in Figure 7.

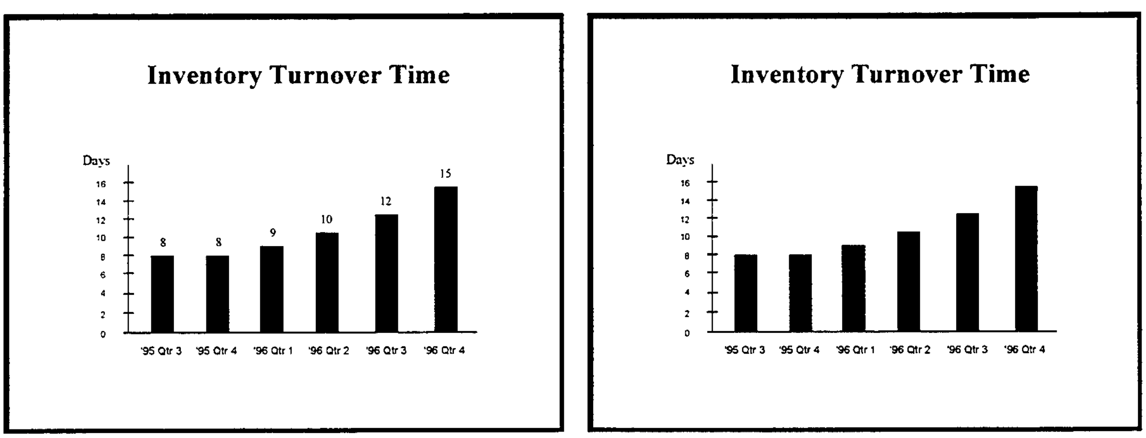


Figure 7. Enhanced Graphics vs. Non-Enhanced Graphics

The results of their study did not show a significant improvement in decision-making due to the enhanced graphics (Corbone and Gorr, 1985: 159). These results suggest that graph presentation is probably beneficial to a point, beyond which further graph enhancements add no additional value to the decision-making process. These enhancements studied by Corbone and Gorr, however, did not include the use of extraneous graphics such as clipart or pictures.

Dickson, DeSanctis, and McBride. Dickson, DeSanctis, and McBride conducted an experiment that focused on the influence of technology on the quality of organizational decision-making (Dickson *et al.*, 1986: 40). Their study compared the effects of graphics and tables on decision-making. The experimental results, which showed no significant differences between graphics and tables, led to the conclusion that "generalized claims of superiority of graphic presentation are unsupported, at least for decision-related activities" (1986: 40). The researchers also concluded that their results suggest the effectiveness of the presentation format is primarily a function of the task at hand, and that conclusions based on "one shot" studies are nothing more that "situationally dependent artifacts" (1986: 40). They also suggested that graphics are more effective as an instrument of persuasion than as a decision support tool (1986: 41). This claim, however, is made with out any empirical evidence, and is stated primarily for the purpose of suggesting further research.

III. METHODOLOGY

An experiment will be used to determine whether extraneous graphics hinder or enhance a viewer's ability to obtain the pertinent information from a professional presentation. The experiment will also be used to determine if extraneous graphics influence the viewer's ability to reach the appropriate decision for the given information, and if the extraneous graphics effect decision confidence. The Methodology section will provide a description of the hypotheses to be tested, the population and subjects, experimental design, apparatus employed, tests conducted, experiment validity, and the method of analysis.

Hypotheses

The hypotheses proposed for this research effort are based on the three research questions presented in Chapter I:

Research Question 1. Do extraneous graphics affect information extraction accuracy?

Null Hypothesis:

H_{01} = Extraneous graphics do not affect information extraction accuracy.

Alternate Hypothesis:

H_{A1} = Extraneous graphics affect information extraction accuracy.

Research Question 2. Do extraneous graphics affect decision accuracy?

Null Hypothesis: H_{02} = Extraneous graphics do not affect decision accuracy.

Alternate Hypothesis: H_{A2} = Extraneous graphics affect decision accuracy.

Research Question 3. Do extraneous graphics affect decision confidence?

Null Hypothesis: H_{03} = Extraneous graphics do not affect decision confidence.

Alternate Hypotheses: H_{A3} = Extraneous graphics affect decision confidence.

The purpose of the first hypothesis is to test if extraneous graphics affect the viewer's ability to accurately obtain the important information from a presentation. It is possible that limited use of extraneous graphics which are related to the subject information may enhance accurate information extraction, whereas extensive use of extraneous graphics are likely to distract the viewer and impede accurate information extraction. The null hypothesis, H_{01}, states that extraneous graphics have no effect on the viewer's ability to accurately obtain the pertinent information from a presentation. The alternate hypothesis, H_{A1}, proposes that extraneous graphics do indeed influence the viewer's ability to accurately extract information.

The purpose of the second hypothesis is to determine if the use of extraneous graphics influences the decision-maker's ability to reach the proper decision indicated by the information provided. As with the first hypothesis, it is likely that some extraneous graphics may help the viewer to make the proper decision, but overuse of extraneous graphics will probably hinder the decision-making process. The null hypothesis, H_{02}, states that there is no effect on the decision-making process regardless of the presence or absence of extraneous graphics. The alternate hypothesis, H_{A2}, proposes that the presence of extraneous graphics will influence the decision-maker's ability to reach the proper decision.

The purpose of the third hypothesis is to determine if the decision-maker's confidence in his or her decision can be influenced by the presence of extraneous graphics. This influence is probably related to H_{A1} since the viewer's decision is likely to be based on the information he or she obtains from the presentation. If decision-makers are uncertain of how well they were able to accurately extract the necessary information, their decision confidence is likely to be affected. Therefore, a correlation analysis will be conducted to determine if a relation exists between the decision-makers' confidence level and their ability to accurately extract information from the presentation.

For the purpose of this experiment, the dependent variable, accuracy (as stated in H_1 and H_2) is defined as the ability of the decision maker to reach the conclusion indicated by the data presented in the graph. Also, for the purpose of this experiment, the second dependent variable is confidence, which (as stated in H_3) is a response by the subject regarding his or her confidence in the conclusion reached.

In addition to these hypotheses, several areas for sensitivity analysis are appropriate. Some previous research has found indications that gender and training (in graphical applications) have an influence on the results obtained (MacKay and Villarreal; DeSanctis). Another potential influential factor for sensitivity analysis is colorblindness. Due to the extensive use of color graphics in the experiment, it is possible that those who are colorblind may perceive and/or interpreted the information differently than those who are not. Analysis was conducted to determine if the influential effects (if any) found in

H_1, H_2, or H_3 differ based on the gender, level of graphics training, or colorblindness of the decision-makers. These three factors are addressed as investigative questions for sensitivity analysis and referred to as hypotheses four, five, and six:

Investigative Question 4. Does gender affect H1, H2, or H3?

Null Hypothesis: H_{04} = Gender does not affect H_1, H_2, or H_3

Alternate Hypothesis: H_{A4} = Gender does affect H_1, H_2, or H_3

Investigative Question 5. Does graphics training affect H1, H2, or H3?

Null Hypothesis: H_{05} = Graphics training does not affect H_1, H_2, or H_3.

Alternate Hypothesis: H_{A5} = Graphics training does affect H_1, H_2, or H_3.

Investigative Question 6. Does colorblindness affect H1, H2, or H3?

Null Hypothesis: H_{06} = Colorblindness does not affect H_1, H_2, or H_3.

Alternate Hypothesis: H_{A6} = Colorblindness does affect H_1, H_2, or H_3.

Population and Subjects

Due to the graphical nature of the media presented to Department of Defense (DoD) decision-makers on a daily basis, the population for this experiment is the set of working professionals (both military and civilian) within the DoD, who make decisions based on briefings, presentations, or other forms of graphical media. These decision-makers very often base their decisions on graphics, and specifically, extraneous graphics.

For example, would project managers be more likely to consider the contract proposal presented on black-and-white, text-only slides, or the contract proposal that includes eye-catching colors and dazzling graphics? Of course we would like to believe that the extraneous graphics do not influence the program manager's decision, but perhaps on a subliminal level, the decision was swayed one way or the other because of the graphics.

The sample populations chosen to represent the subset of the DoD decision-makers consisted of AFIT Professional Continuing Education (PCE) students, AFIT graduate students, and AFIT professors. PCE students were chosen to participate in the experiment because they typically consist of a representative mix of civilian and military, as well as a mixture of age, race, gender, rank/grade, experience, and education. The majority of test subjects consisted of PCE students. The AFIT graduate students and instructors provide a representative mix of company grade officers, field grade officers, and middle to senior level civilians.

The PCE classes, graduate students, and instructors selected for this study were dependent upon course director and instructor approval. No specific classes or individuals were targeted for use in the sample. By not targeting specific courses or individuals, the sample includes a broader base of the population. Although convenience sampling is considered by some as an unreliable source of data, the power of this test will come from the large sample sizes that were used. Additionally, the use of students for exploratory research has been employed in most previous research.

> [F]or activities of the type our experiments require, there is no reason to believe that students would perform any differently than managers. Additionally, for this kind of work, one needs the power in tests that can be obtained by the large numbers available in student subjects. (Dickson *et al.*, 1986: 46)

Experimental Design

The design of the experiment is a 3 X 1 factorial. The single factor is the application of extraneous graphics. Extraneous graphics are defined as any non-essential images, drawings, pictures, or logos added to a presentation. The three levels for the single factor are three different treatments of extraneous graphics. The first treatment has no extraneous graphics. The second treatment contains a medium amount of extraneous graphics, and the third treatment contains a high amount of extraneous graphics (See the Apparatus section of this chapter for examples and explanations of these three treatment levels).

The response variables used to test the Hypotheses H_1, H_2, and H_3 are information extraction accuracy, decision accuracy, and decision confidence. For information extraction accuracy a correct response was coded as a "0" and an incorrect response as a "1." The format of the test questions requested either an "agree" or "disagree" response from the test subjects. The subjects indicated their response on the provided answer sheet, and their answers were assigned "0" if correct and a "1" if incorrect. Appendix C contains the response data received from the test subjects. For decision accuracy and decision confidence the test subjects were asked to provide a scaled response with a "1" indicating "strongly disagree" and a "7" indicating "strongly agree," as shown in Figure 8. The participant's responses were simply recorded in the original numeric format.

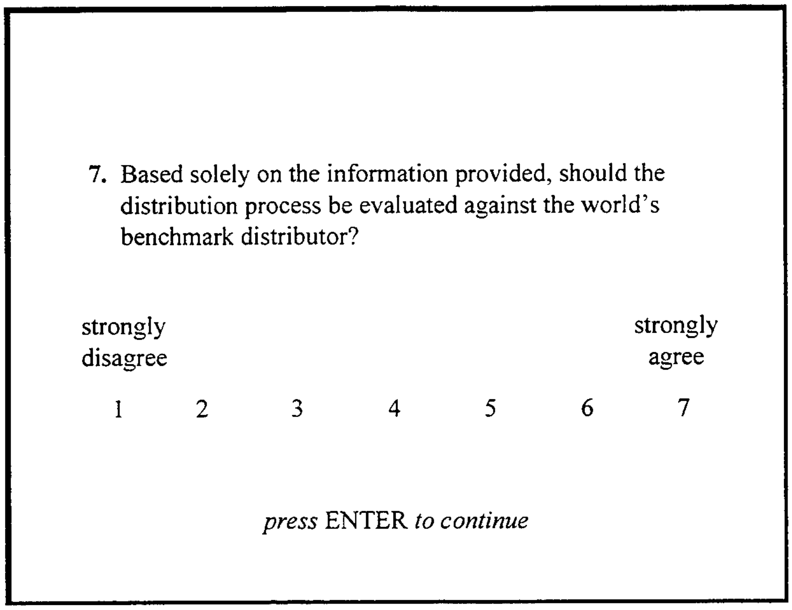


Figure 8. Presentation Slide containing the Decision Question

Apparatus

A PowerPoint® presentation containing six informational slides followed by fourteen question slides make up the experimental item. Appendices A and B contains the complete experimental items. There are three versions of the presentation (one for each of three treatment levels), in which only the first six informational slides are affected by the differing treatments. The fourteen question slides remain the same for all three treatments. The presentations were viewed using the Slide Show function in PowerPoint, and displayed on a personal computer (PC) monitor. The test subjects were randomly assigned to one of the three versions, and individually viewed the presentation on a PC monitor. An instructional sheet, answer sheet, and demographics questionnaire were also provided for each participant. The answers to the questions within the presentation were recorded by the subject on the provided answer sheet.

The decision problem chosen for this experiment involves benchmarking. Benchmarking refers to the process of comparing an organization with a better organization (preferably the "world's best") to evaluate the first organization's performance and identify problem areas. This same decision problem (with minor modifications) was used by Peeler, and was chosen so that the results of this experiment could be compared with his results. Peeler's study analyzed another form of extraneous graphics referred to as background graphics.

The decision problem consists of six informational slides that provide the necessary information to reach an appropriate decision. To address the first hypothesis (regarding information extraction accuracy), the viewers are asked one information extraction question for each of the six slides - for a total of six questions. Then to address the second hypothesis (decision accuracy), the viewers are asked to make a decision regarding whether or not a benchmarking effort is appropriate for the given situation. The provided financial and performance trends clearly indicate which decision alternative is appropriate. Following the decision, the viewers are asked to rate the level of confidence they have in their decision (for the third hypothesis). Finally, a series of opinion questions address the personal preferences of the test subjects regarding the use of extraneous graphics.

The slides presented in Figures 9 through 14 are those presented to the decision-makers. The slides in these figures represent the "no extraneous graphics" condition. Two other versions, one with a "medium" and one with a "high" amount of extraneous graphics, will be discussed in the subsequent paragraphs. Each test subject in the experiment is cast as the Chief Executive Officer (CEO) of a large manufacturing firm that produces and distributes ball bearings. It has come the management's attention that something is wrong with the company's distribution process, and no one within the company can identify the problem. Based solely on the information provided in the presentation, which covers the performance of the company over the past six quarters (1 ½ years), the viewer must decide whether or not to consult the world's most efficient distributor for assistance in evaluating the company's ailing distribution process.

All of the slides are constructed in order to clearly indicate that the company should indeed benchmark its distribution process. The first slide, shown in Figure 9, represents the firm's distribution costs, which have significantly increased during the past six quarters. The distribution costs for the firm's closest competitor has only increased slightly during this same time period. As a result, the competitors distribution costs are now lower than the subject firm's. The information extraction question for this slide asks if the firm's distribution cost per item has increased more than the competitor's distribution cost per item. The correct response to this question is "agree." Obviously the firm's distribution costs have risen dramatically in recent months and some action is necessary to identify the cause of this problem.

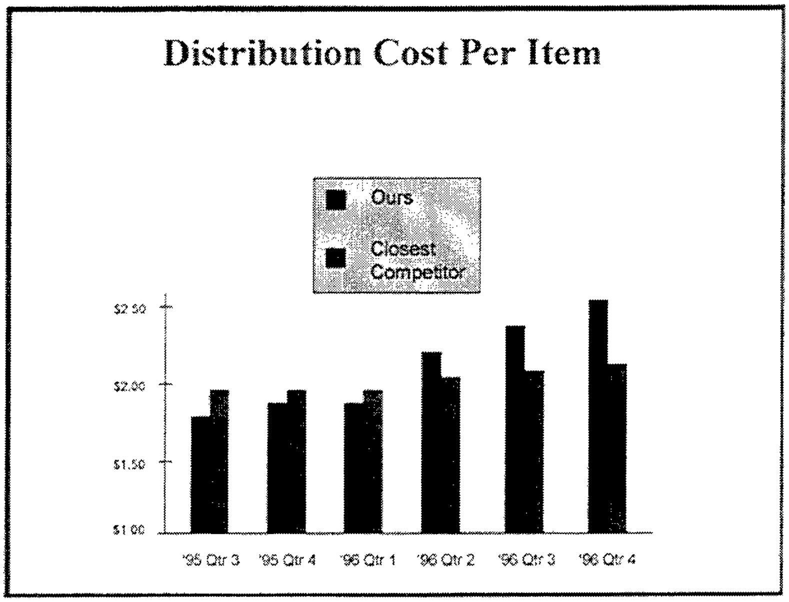


Figure 9. First Slide of the Benchmarking Decision-Problem

Figure 10 shows the financial strength of the firm compared to its closest competitor. The graph shows that the firm's stock prices have steadily declined during this time period, while the competitor's stock has gained in value. The information extraction question asks if the firm's stock prices have declined over the period shown. The correct answer to this question is "agree," since the trend shows a definitive decrease in the firm's stock price per share. The firm's market share is also decreasing, as shown in Figure 11. The related question asks if the firm's market share has been declining over the past three quarters, to which the correct answer is "agree."

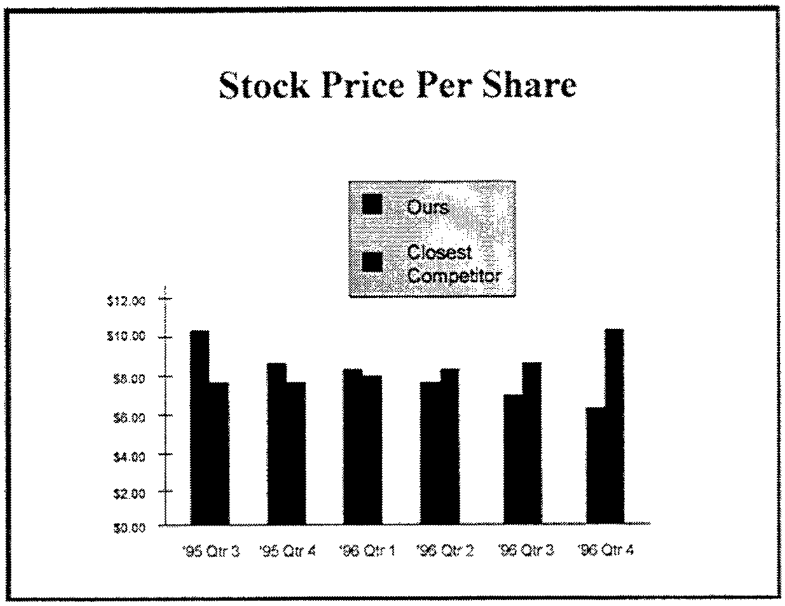


Figure 10. Second Slide of the Decision-Problem

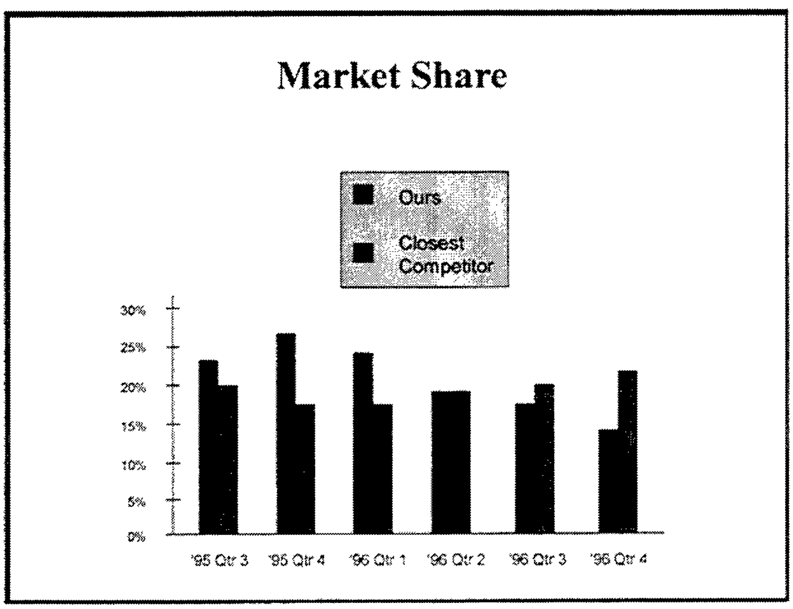


Figure 11. Third Slide of the Decision-Problem

More information in favor of the benchmarking decision is presented in Figure 12, which shows the firm's cycle time for order processing. The time required to process an order has dramatically increased over the past three quarters. The information extraction question related to this siide asks if the cycle time has decreased over the past three quarters. The correct response to this question is "disagree," since the cycle time has actually doubled during this time period.

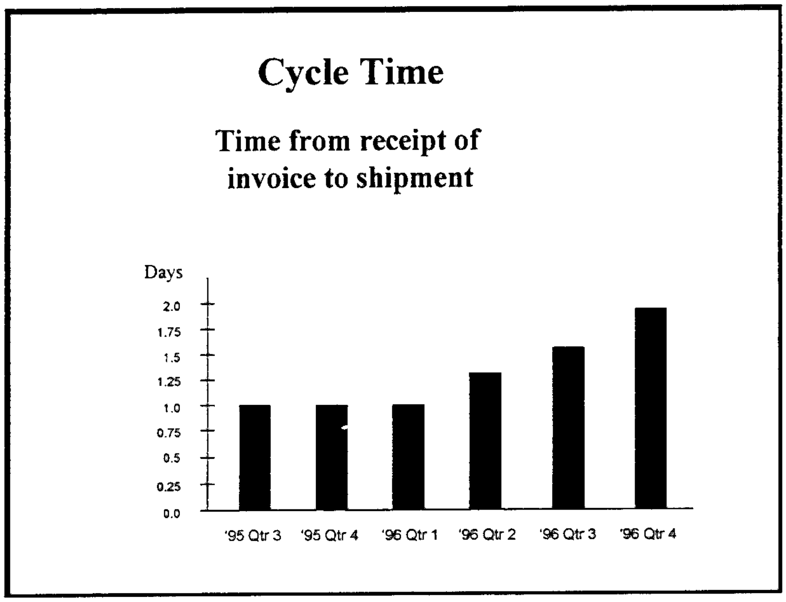


Figure 12. Fourth Slide of the Decision-Problem

While the cycle time has increased, the time required to go through the inventory has also gone up. Figure 13 shows that the inventory turn over time has more than doubled over the past six quarters. Therefore the correct answer to the question, which asks if the time needed to go through inventory (turn over time) is declining, is "disagree."

The final informational slide conveys the message that the firm's internal problems are being noticed by their customers. Figure 14 shows that the customer satisfaction with the firm's distribution system and delivery time has dramatically fell from a respectable “8” to an embarrassing “3” (approximately). The related question asks if customer satisfaction has declined over the period shown. The correct response to this question is definitely “agree.”

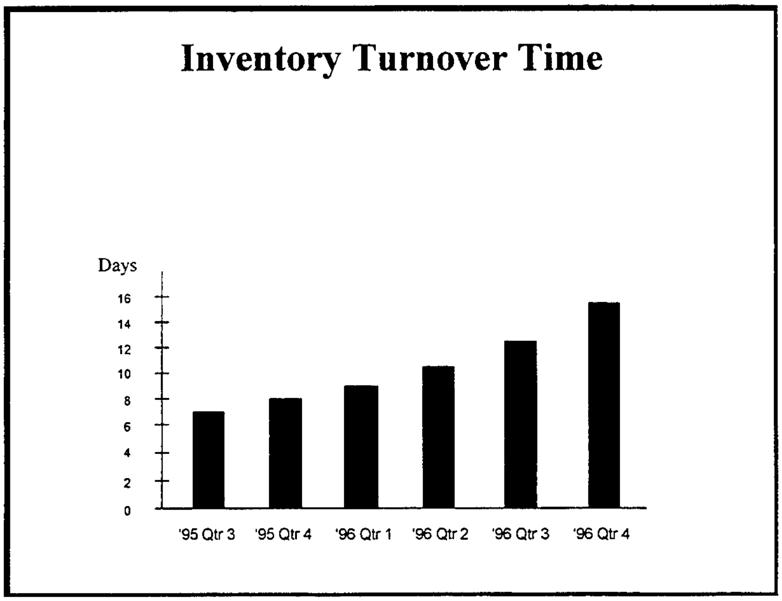


Figure 13. Fifth Slide of the Decision-Problem

Taken as a whole, these trends certainly indicate that the firm needs to take action in order to locate and correct the problems causing this rapid decline in performance. The use of benchmarking is indicated as the most likely means of identifying the problem since no one within the organization has be able to do so. Using an external benchmark, who is an expert in the distribution process, seems to be the appropriate decision.

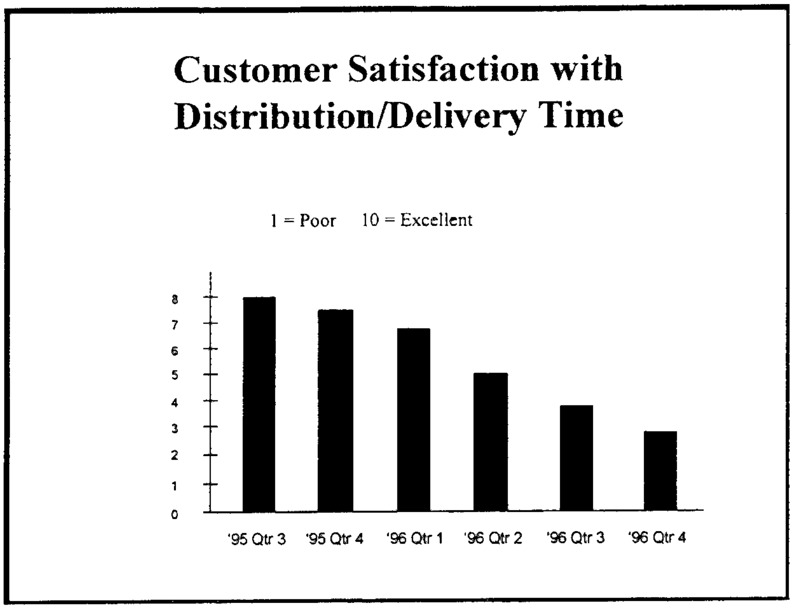


Figure 14. Sixth Slide of the Decision-Problem

Three versions of the decision problem, described above, were used to test the effects of extraneous graphics. The first version has no extraneous graphics and is exactly like the slides shown in Figures 9 through 14. The second version includes a "medium" amount of extraneous graphics incorporated into the slides. For the purpose of this research effort, medium refers to approximately 2 or 3 clipart images, for a total of about 14% of the slide's surface area. The third version includes a "high" amount of extraneous graphics, which constitutes about 3 to 5 clipart images (depending on their sizes), totaling of approximately 23% of the slide's surface area. Figures 15 and 16 illustrate the Medium and High versions of one slide in the decision-problem. The original information, as shown in Figure 13, has not been moved or altered in any way in order to maintain parallel construction between all three treatments. The complete presentations for the Medium and High treatment versions can be found in Appendix B.

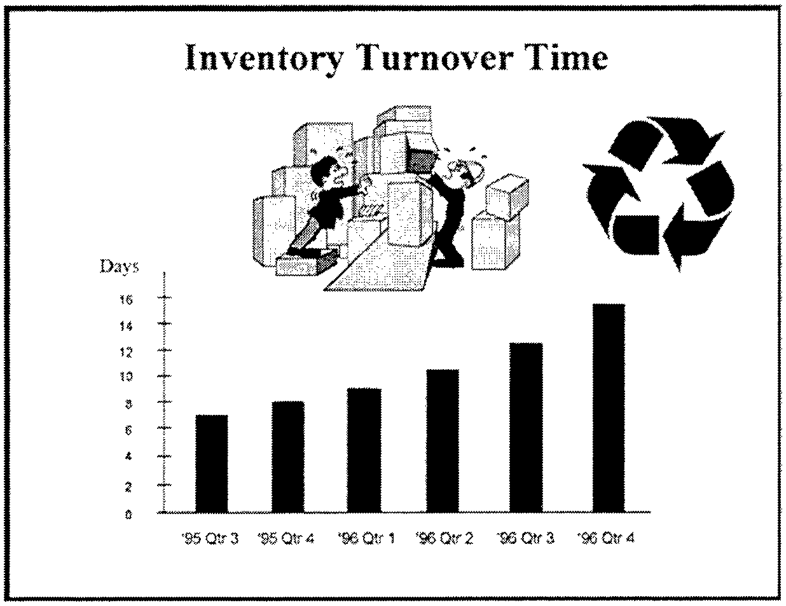


Figure 15. Medium Amount of Extraneous Graphics

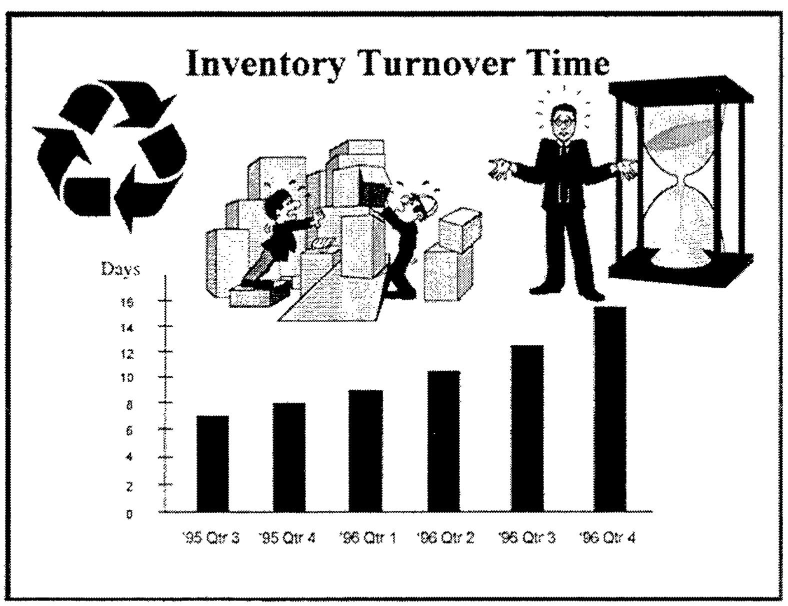


Figure 16. High Amount of Extraneous Graphics

The clipart images used in these presentations were obtained from either the PowerPoint® 7.0 clipart library, or from Swift Platinum's Flexomatic Clip Art® software package. The specific images chosen for use in the experiment were deemed to be reasonably related to the subject matter of the particular slide on which they were placed. The objective is to imitate the common practice of including clipart that is somewhat relevant to the presentation, but not necessary to convey the pertinent information - hence the term "extraneous." The clipart used in this experiment is not intended to provide additional information, and is arranged in such a manner as to avoid obscuring or deleting any of the original information.

Following the decision problem slides, the test subjects are asked a series of questions regarding their opinion of the extraneous graphics (or lack thereof) used in the presentations. Figure 17 shows the first three opinion questions. Questions 9 and 10 deal with the viewer's opinion regarding the appropriateness of the extraneous graphics they encountered. [The term "support graphics" is used in these questions to avoid any negative connotations that might be associated with the word "extraneous."] Question 11 asks the viewers to assess their own perception of the effects extraneous graphics may have had on their decision-making process.

9. Do you feel that the presentation contained the appropriate amount of support graphics, such as clip-art, logos, icons, etc., (not including the informational bar charts) for a CEO-level briefing?

strongly disagree						strongly agree
1	2	3	4	5	6	7

10. Was there too much or too little use of support graphics?

too little						too much
1	2	3	4	5	6	7

press ENTER *to continue*

11. What was the impact of the support graphics (clip-art) in the presentation on your ability to understand the information and make a decision?

very detrimental						very helpful
1	2	3	4	5	6	7

press ENTER *to continue*

Figure 17. Opinion Questions that follow the Decision-Problem

The last three opinion questions ask the viewers for their preference regarding which amount of extraneous graphics (None, Medium, or High) they consider to convey a more professional appearance. To do so, three direct comparison slides like the one shown in Figure 18 were used. The example in Figure 18 shows the comparison between the None and Medium conditions. The two other comparison slides not shown, compare the High and Medium conditions, and the None and High conditions.

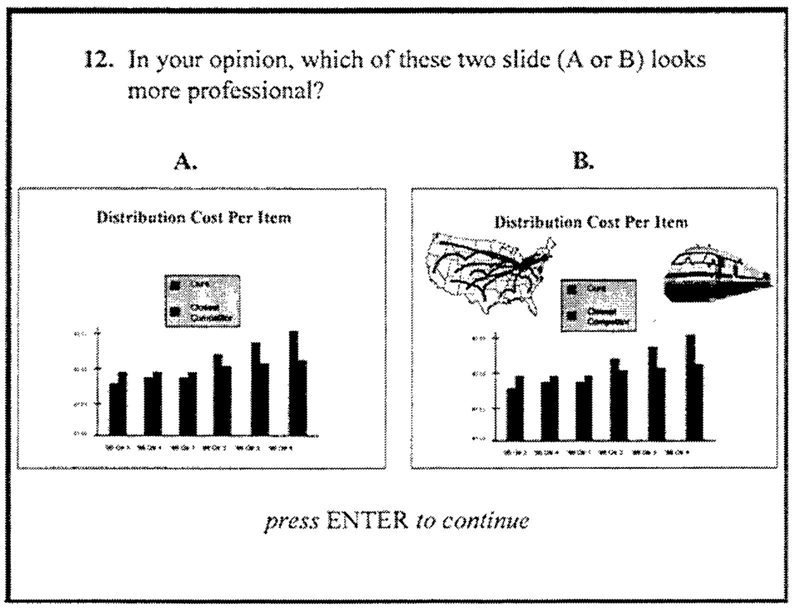


Figure 18. Comparison Slide for Viewer Preference

Following the completion of the last opinion question slide, the test participants are then asked to complete a demographics questionnaire in order to obtain the necessary information for sensitivity analysis (Hypotheses 4, 5, and 6). The entire task (including reading the instructions and completing the questionnaire) requires approximately 15 minutes to complete. The first six slides, which contain the information for the decision-problem, are programmed to appear for 15 seconds each, and will automatically advance at the end of the 15 second period. During this time interval, the participants' only requirement is to observe the slide. This time interval was chosen to allow enough time for adequate viewing without providing too much time for over-analysis. Furthermore, decision-makers in the DoD population are often required to make quick decisions without the benefit of reflection and excess thought. The test questions which follow the six informational slides are also displayed on the computer screen. There was no time limit for answering the questions at the end of the presentation, and the participants were able to manually advance to the next questions after recording their answer on the provided answer sheet.

Tests Conducted

Pilot Tests. Prior to administration of the experimental apparatus, pilot testing was conducted using AFIT professors and graduate students. A total of 23 subjects participated in the pilot tests. The first iteration of the experiment was tested using 12 subjects, and the second iteration of the test used the remaining 11 subjects. The size of the subject groups in both of these tests was kept relatively small in order to facilitate comprehensive feedback from the participants. The first iteration of the pilot test included two possible experimental designs. One design included the information and the question regarding that information on the same slide. In this format the viewer simultaneously had access to both the question and the information necessary to answer the question. The other design placed the information and questions on separate slides.

In this format the subjects were shown all six informational slides first, followed by additional slides containing the questions. Both designs used programmed time intervals to automatically advance every slide, including the question slides.

The second design (separate slides for the questions) was the preliminary favorite since having the questions on the same slides as the information may cause the viewer to focus strictly on the information necessary to answer the question, and thus, ignoring (or avoiding) the potential influences of the extraneous graphics. If the viewers does not know what they will be questioned on, they will focus on whatever naturally attracts their attention - which may or may not be the pertinent information.

During the feedback sessions, the participants were asked to provide their opinions and observations regarding the validity of each design. Of particular concern for the second design, was the ability of the subjects to recall the information necessary to answer the questions, *after* having seen all of the informational slides. Despite our concerns, the test subjects unanimously concluded that the second design seemed to be more realistic and more valid. Feedback was also sought regarding the programmed time interval allotted to view each slide and the clarity of the instructions.

Of the subjects who tested the second design, most supported the use of timed advancement of the informational slides, but did not see a need for limiting the time allotted to answer the questions. Since the questions were separate from the information, pausing for a few seconds to ponder a question would not allow the subject to see any more information, as would be the case for the first design (assuming manual advancement). Regarding clarity of the instructions, most subjects felt that more "on-screen" instructions would be beneficial. Primarily, the subjects wanted a little more direction as to what they were supposed to do at the beginning and end of the presentation. Feedback was also received related to the wording of the questions, additional questions to include, and general appearance/format issues.

Based on the feedback from the first test run, the experiment design using separate slides for the information and questions was selected. The feedback also led to other modifications such as automatic, timed slide advancement for the informational slides only. The question slides were changed to a manual advance format. The wording of a few questions was modified, and the direct comparison questions (illustrated in Figure 18) were added. Finally, a few additional instructions were included and the existing instructions received some clarifications.

Following the first administration of the pilot test and the subsequent modifications, a second iteration was tested. The second iteration received very positive feedback and the experimental design was determined to be "mission capable."

The Experiment. At the start of the experiment, each participant was seated at a PC with a title slide visible on the monitor. The title slide informed the viewer to first read the provided instructions and scenario (but not the questionnaire), and then press the Enter Key when ready to view the presentation. Following the title slide, the six informational slides that make up the experimental presentation advanced automatically after providing 15 seconds to view each slide. The seventh slide (not counting the title slide) contained the first question. The remaining question slides were manually advanced by the participant after recording his or her response on the provided answer sheet. When answering the questions, the subjects were not permitted to review the informational slides. The final slide informed the participant that the automated portion of the experiment had been concluded and asked them to complete the demographics questionnaire. Complete copies of the experimental apparatus can be found in Appendices A and B.

Experiment Validity

There are many potential threats to the internal and external validity of any experiment. Internal validity questions whether the demonstrated experimental relationship truly implies cause, and External validity questions whether a casual relationship can be generalized across the population.

There are seven major threats to internal validity:

> 1. *History.* During the course of an experiment, events may occur that confound the effects.
>
> 2. *Maturation.* Changes may occur, which are not specific to any event, that are a function of the passage of time.
>
> 3. *Testing.* The process of taking a test can influence performance on a second test.
>
> 4. *Instrumentation.* Changes may occur in the measuring instrument or the observer.
>
> 5. *Selection.* Differential selection of subjects leads to biased comparisons.
>
> 6. *Statistical Regression.* This factor is especially predominant when groups have been selected on the bases of their extreme scores.
>
> 7. *Experimental Mortality.* This occurs when the composition of the subject groups changes during the experiment (Cooper and Emory, 1995: 358-359).

For this thesis design, history is not a factor, and maturation and mortality are very unlikely because the test is administered in one sitting consisting of approximately 10 to 15 minutes. Also due to the single sitting format, testing influences are not a threat. Had the experimental questions been placed on the same slides with the relevant information, a "within-test" testing influence may have occurred. After observing the test format on the first few slides, the subjects would have learned what to look for on the subsequent slides. This potential threat was avoided by placing the questions on separate slides at the end of the presentation.

Instrumentation is not a problem, because the format, structure, and content of the measuring instrument were held constant. The automated format of the presentation should preclude any changes from occurring. When the computer laboratories were used for test administration, the experiment was loaded from original floppy diskettes (controlled by the author) each testing session, to avoid potential problems such as tampering or corruption of network or locally stored files.

Potential threats associated with selection or statistical regression are mostly nullified by random selection and assignment of test subjects (Cooper and Emory, 1995: 359). Randomization was accomplished on two levels for this experiment. On the first level, the members of a continuing education course, which is available to all DoD personnel, is a reasonably representative random sample of the DoD population. Likewise, AFIT graduates students also represent a fairly random sample of company grade officers and their equivalent civilian counterparts. On the second level, the test subjects were randomly assigned to one of three experimental groups. Therefore, randomization on two levels should preclude any selection or statistical regression threats.

There are two primary threats to external validity, as well as a few other reactive factors:

> 1. *Reactivity of Testing on Experimental Treatment.* Pretests or other preliminary exposure to the experiment, to include multiple treatments, sensitizes the subjects to the experimental stimulus.
>
> 2. *Interaction of Selection and Experimental Treatment.* The process for test subject selection is a threat if the population from which the subjects were selected does not match the population for generalization.
>
> 3. Other Reactive Factors such as *artificial experimental settings* and *knowledge of experiment participation* are threats to external validity. (Cooper and Emory, 1995: 360-361)

Reactivity effects are mostly mitigated by the design of this experiment. The single-step administration of the test avoids compromising the participants sensitivity to the experiment. Pretests or other preliminary testing activities were not conducted in order to prevent premature exposure to the test item. Any individual with prior knowledge of the experiment was precluded from the sample.

Selection biases, as discussed under internal validity, were mitigated as best as possible using two levels of randomization. The population from which the test subjects were drawn is reasonable representation of the DoD decision-making population. Had a much broader population such as the "general public" been claimed, then the reasonableness of the sample representation would have been reduced.

The experimental arrangements are somewhat artificial due to the academic environment, and a minor threat to external validity. However, the potential biasing effects of this setting were somewhat mitigated by the timing and location of the testing for PCE students. Since the test required the use of a PC, the testing was administered to PCE courses that are taught in the computer lab. This precluded any requirement for mass movement from a classroom to an computer lab. The students were, therefore, tested in their "familiar" environment, using a standard software package (PowerPoint), which is commonly used in the DoD for decision-style presentations.

Method of Analysis

The first step in determining the method of analysis is to decide between parametric or non-parametric tests. The use of parametric tests requires meeting four criteria:

1. The observations must be independent
2. The observations should be drawn from normally distributed populations
3. The populations should have equal variances
4. The measurement scales should be at least interval so that arithmetic operations can used. (Cooper and Emory, 1995: 443)

Some of the data collected for this research effort, however, do not meet these four criteria. Since the responses for the first six test questions were categorized as either a "0" or "1," the resulting data is non-continuous, ordinal data. This non-continuous data is not normally distributed, and the requirement for equal variances is, likewise, not satisfied.

Therefore, the Kruskal-Wallis non-parametric test will be employed to analyze the data. The three criteria for using the Kruskal-Wallis are:

1. All samples are random samples from their respective populations.

2. In addition to independence within each sample, there is a mutual independence among the various samples.

3. The measurement scale is at least ordinal. (Conover, 1971: 230)

Because the ordinal data collected during this experiment is from independent random samples, the Kruskal-Wallis test is applicable. Another benefit of the K-W test is that it does not require equal sample sizes, which is the case for this experiment.

The Kruskal-Wallis test is designed to be sensitive against differences in means among samples. The K-W test ranks the observations from 1 to N, where N is the largest random variable. When the null hypothesis is true (sample means are equal), the N observations all come from the same distribution, in which case all possible assignments of the ranks are equally likely. If null hypothesis is false, then some samples will consist mostly of observations having small ranks while others will consists mostly of observations having large ranks. More specifically, if R_{ij} denotes the rank of an observation among N observations, while R_i denotes the total of the ranks in the ith sample, and $\overline{R_i}$ denotes the average of the ranks in the ith sample, then when the null hypothesis is true (E denotes expected value):

$$E(R_{ij}) = \frac{N+1}{2} \quad \text{and} \quad E(\overline{R_i}) = \frac{1}{J_i}\sum_j E(R_{ij}) = \frac{N+1}{2} \qquad \text{(Devore, 1995: 652)}$$

The K-W test statistic is a measure of the extent to which the $\overline{R}_i$'s deviate from their common expected value $E(R_{ij})$. The null hypothesis is rejected if the computed value of the K-W statistic indicates too large of a discrepancy between the observed and expected rank averages. The K-W test statistic (K) is defined as:

$$K = \frac{12}{N(N+1)} \sum_{j=1}^{I} Ji \left(\overline{R_i} \cdot - \frac{N+1}{2} \right)^2 = \frac{12}{N(N+1)} \sum_{i=1}^{I} \frac{R_{i\cdot}^2}{J_i} - 3(N+1)$$

(Devore, 1995: 652)

The purpose of the analysis is to determine if there is a statistically significant difference between the responses for each of the three treatment levels. If a difference is shown to exist, then the null hypothesis will have to be rejected, thus, accepting the alternate hypothesis. Rejecting the null hypothesis, therefore, indicates that extraneous graphics do indeed effect the variable in question (extraction accuracy, decision accuracy, or decision confidence)

If no statistically significant difference is found then the null hypothesis is not rejected. This condition would indicate that no significant effects of extraneous graphics could be shown for the respective variable.

IV. ANALYSIS AND FINDINGS

An experiment was used to test the possible effects of extraneous graphics on information extraction accuracy, decision-accuracy, and decision confidence. The experiment consisted of a decision problem in the form of a PowerPoint® presentation. The test subjects responded to a series of questions regarding the decision problem proposed in the presentation. The data obtained was in the form of the test subjects' responses, and was used to perform analysis with respect to the proposed hypotheses. The purpose of this chapter is to present the results of the analysis and discuss the interpretation of the findings.

The decision problem was used to test the following three hypotheses (the null hypotheses are listed):

H_{01}: Extraneous graphics do not affect information extraction accuracy.

H_{02}: Extraneous graphics do not affect decision accuracy.

H_{03}: Extraneous graphics do not affect decision confidence.

In addition to the above hypotheses, three areas for sensitivity analysis were identified as potential influences on the first three hypotheses. These areas were treated as investigative questions and are referred to as hypotheses four, five, and six:

H_{04}: Gender does not affect H_1, H_2, or H_3.

H_{05}: Graphics training does not affect H_1, H_2, or H_3.

H_{06}: Colorblindness does not affect H_1, H_2, or H_3.

Other data, not directly related to the proposed hypotheses, were also collected. This data were analyzed to provide further insights into the preferences and perceptions of DoD professionals regarding the use of extraneous graphics.

Test Subjects

The test subjects who participated in the thesis experiment were either enrolled in AFIT Professional Continuing Education (PCE) courses or AFIT graduate degree programs, or were AFIT faculty members. A total of 101 test subjects participated in the experiment. Each of these subjects was randomly assigned to one of the three experimental treatment groups. Thirty-one subjects viewed the treatment condition with no extraneous graphics, 36 viewed the test condition with a medium amount of extraneous graphics, and 34 subjects viewed the test condition with a high amount of extraneous graphics. The different sample sizes for each test condition were due to the random assignment of test subjects. All of these test subjects were asked to complete a questionnaire to provide demographics information such as gender, education level, and graphics training. This information was used to verify the random assignment of test subjects and to validate the generalizability of the test sample to the population (DoD decision-makers). This information was also used to conduct sensitivity analysis regarding gender, graphics training, and colorblindness. Table 2 provides a summary of the demographics information obtained from the test subjects.

Table 2. Demographics Information

Demographic	No Extraneous Graphics	Medium Extraneous Graphics	High Extraneous Graphics	Total
Sample size	31	36	34	101
Females	10	11	15	36
Males	21	25	19	65
Doctorate	0	1	1	2
Masters	5	7	4	16
Graduate courses	17	15	9	41
Baccalaureate	5	7	11	23
Associates	3	3	2	8
College courses	1	3	6	10
High School	0	0	1	1
Not Colorblind	31	33	33	97
Colorblind	0	3	1	4
No Training	9	11	15	35
Training	22	25	19	66

The distribution of gender, education levels, colorblindness, and graphics training indicate random assignment of test subjects. As would be expected from random assignment, the relative numbers of test subjects in each category are not equally distributed across the three treatments, but are basically similar.

The education level of the subjects varied from High School diplomas (1%) to Doctoral degrees (2%), with the largest percentage of participants (41%) having completed "some graduate courses." Exact figures identifying the relative percentages of each education level within the DoD population were not available. However, the test sample distribution of education levels seems to approximate the population distribution of DoD decision-makers, which consists of both military enlisted and officers, as well as civilian government professionals.

One area identified for sensitivity analysis is the potentially influencing effect of gender on extraction accuracy, decision accuracy, or decision confidence. Of the 101 test subjects, 64% were male and 36% were female. The results of the sensitivity analysis pertaining to gender performance, are discussed in the following section entitled "Experimental Results."

Another area identified for sensitivity analysis is the potential effects of graphical training on extraction accuracy, decision accuracy, or decision confidence. The questionnaires showed that 65% of the subjects had received either formal or informal training with graphics/presentation construction or interpretation. The performance of this test group was compared to that of the other 35%, who had not received any formal or informal training in this area. The results of this analysis are discussed in the "Experimental Results" section.

The third area identified for sensitivity analysis is the potential effects of colorblindness. Due to the extensive use of colored graphics in the experiment, it is possible that those who are colorblind may have interpreted the information differently than those who are not. However, the number and distribution of colorblind test subjects was too small to conduct meaningful statistical analysis. A total of four subjects claimed to be colorblind, three of which viewed the medium extraneous graphics treatment, one saw the high treatment, and none viewed the treatment without extraneous graphics. It is possible that more of the test subjects are actually colorblind, but do not know of their condition. However, the time limitations and scope of this experiment did not allow for administering colorblindness tests to all subjects.

Experimental Results

The data were analyzed using quantitative statistical analysis. Two forms of data were obtained from the experiment. The first form is ordinal data, which was acquired from the first six test questions, and used to test for information extraction accuracy (Hypothesis H_1). The last three test questions (questions 12 - 14) also provided ordinal data, but statistical analysis was not required for these test questions. The second form is interval data, which was acquired from the remaining test questions (questions 7 - 11). The responses to question 7 were used to test for decision accuracy (H_2), and the responses to question 8 were used to test for decision confidence (H_3).

For the first six test questions, the subject's responses were either correct (recorded as a "0") or incorrect (recorded as a "1"). Table 3 provides a summary of this ordinal response data compared across the three treatment levels. The data in the table indicates the number of subjects' responses that were correct and incorrect for the first six questions in each treatment level. The 31 test subjects who viewed the experimental treatment with no extraneous graphics answered a total of 186 questions pertaining to extraction accuracy (31 X 6 = 186), of which 170 responses were correct and 16 responses were incorrect. When analyzing the first six test questions as a whole, the 31 subjects who missed 16 questions averaged 0.52 incorrect responses per person. Therefore, each test subject in the "no extraneous graphics" group missed about one half of a question out of the first six.

The 36 test subjects who saw the experimental treatment with a medium amount of extraneous graphics answered 194 questions correctly and 22 incorrectly, for an average of 0.61 incorrect responses per person. The 34 subjects who viewed the high extraneous graphics treatment answered 169 question correctly and missed 35, for an average of 1.03 incorrect responses per person.

Table 3. Summary of Data and Results for Information Extraction Accuracy.

	No Extraneous Graphics	Medium Extraneous Graphics	High Extraneous Graphics	K-W Test Statistic	P-Value
Sample Size	31	36	34		
Total Responses	186	216	204		
Correct Responses	170	194	169	7.8045	0.0202
Incorrect Responses	16	22	35		
Incorrect Responses per person	0.52	0.61	1.03		

Table 3 also contains the results of statistical analysis employing the Kruskal-Wallis nonparametric analysis of variance. Both the Kruskal-Wallis test statistic and the corresponding P-value are shown. Additionally, the Chi-Square test for heterogeneity was conducted using this same data to verify the results of the Kruskal-Wallis test. The Chi-Square test resulted in a P-value of 0.0201, which is almost identical to the P-value found using the K-W test.

These results indicate that, for confidence levels up to 98%, the differences between experimental treatments are indeed statistically significant for the total responses from the first six test questions. Therefore, for α values greater than 0.02 (i.e. confidence levels less than 98%) this finding appears to justify rejection of the null hypothesis H_{01}: extraneous graphics do not affect information extraction accuracy. Conducting the same type of analysis for each of the first six test questions individually, do not reveal that any particular question, or questions, were individually responsible for the total effect.

The general trend appears to indicate more incorrect responses as the amount of extraneous graphics increase. Further analysis was conducted to confirm this observation. Figure 19 illustrates the actual mean values (with 95% confidence intervals) of the data for each of the three test conditions. The leftmost bar represents no extraneous graphics, the center is for the medium treatment, and the rightmost bar is the high treatment. This Figure shows the *actual* mean values and not the average rank values ($\overline{R_i}$) which were used for the K-W test. Therefore, the overlapping confidence intervals shown in this figure do not contradict the K-W and Chi-Squared test results, which indicated a statistical difference between the three treatment levels. Note that these mean values are from data sets containing only ones and zeros, most of which are zeros (i.e. correct responses). The sample with a larger mean value contains more ones in the data, and therefore, more incorrect responses. It appears from the graph that the amount of incorrect responses increases with the amount of extraneous graphics. Although the separation between the means of the none and medium treatments appears to be very small, the high treatment seems to be the most different.

To verify these observations, statistical analysis using the K-W test, was used to determine which treatments were causing the significant variance, shown as the low P-value in Table 3. Directly comparing each of the three possible combinations of treatment levels was the technique used to conduct this analysis. The results of this analysis are shown in Table 4. As expected, the results indicate no statistically significant difference between the none and medium test conditions, but the differences between the none and high conditions, as well as between the medium and high conditions are significant for confidence levels up to 96%.

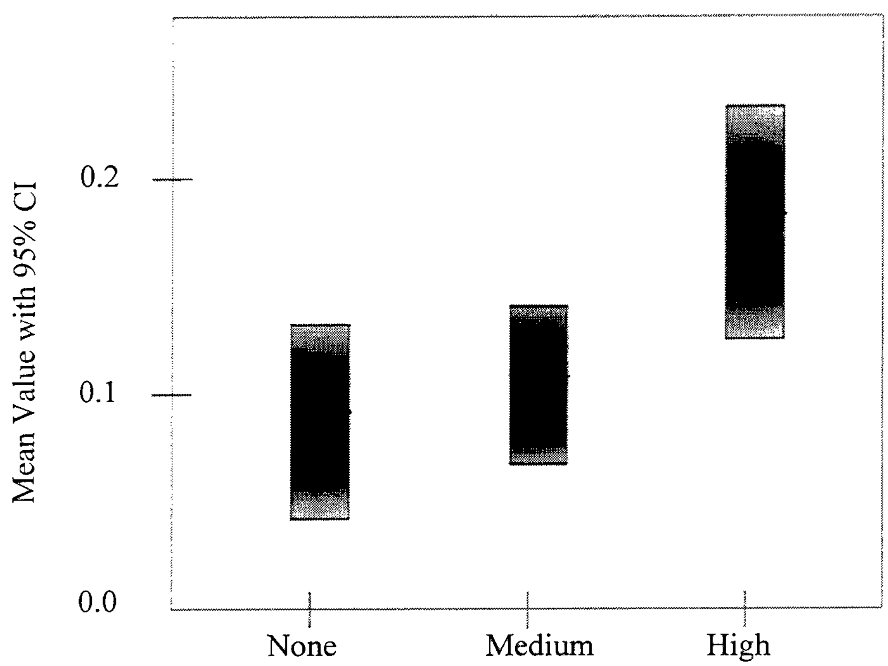


Figure 19. Comparison of Means for Test Questions 1 through 6 (in total)

Table 4. Direct Comparison of Test Conditions

Comparison Pair	K-W Statistic	P-Value
None and Medium	0.2919	0.5890
Medium and High	4.3370	0.0373
None and High	6.2479	0.0124

Therefore, the statistical analysis of the data from the first six test questions, indicates that a medium amount of extraneous graphics has no significant effect on information extraction accuracy. However, a high amount of extraneous graphics distracts the viewer and has a negative effect on information extraction accuracy.

To test H_1 for sensitivity to gender, the information extraction performance of male test subjects was compared to that of female subjects. Table 5 summarizes the results of the gender sensitivity analysis. The number of correct and incorrect responses for each test condition are broken down according to gender (M = male and F = female), and the resulting P-value from the Kruskal-Wallis test within each test condition are provided. The P-value between treatments within each gender group is also shown.

Table 5. Sensitivity Analysis: Gender vs. Information Extraction Accuracy

	No Extraneous Graphics		Medium Extraneous Graphics		High Extraneous Graphics		Between Treatments	
	M	F	M	F	M	F	M	F
Sample Size	21	10	25	11	19	15	65	36
Total Responses	126	60	150	66	114	90		
Correct Responses	117	53	136	58	102	67		
Incorrect Responses	9	7	14	8	12	23		
Incorrect Responses per person	0.43	0.70	0.58	0.73	0.63	1.53		
P-Value	0.3050		0.5336		0.0048		0.6457	0.0347

Examination of the resultant P-values reveals that male and female subjects demonstrated similar competence in information extraction for both the none and medium extraneous graphics treatments. However, the extremely low P-value in the high extraneous graphics treatment indicates a significant difference between the information extraction accuracy of males and females for confidence levels up to 99.5%. This finding indicates that women are much more likely to be distracted by high amounts of extraneous graphics, and subsequently less accurate in their information extraction. It is interesting to note that numbers of incorrect responses per test subject are higher for female participants in all three test conditions, but the difference is statistically significant in the case of high extraneous graphics. Females, who make up 44% of the test subjects in the high extraneous graphics condition, accounted for 68% of the incorrect responses. The ratio of incorrect responses per person increased with increasing levels of extraneous graphics for both men and women, but the effect seems to be more dramatic for women. In fact, female test subjects are responsible for the majority of the differences across all three test conditions, which led to the rejection of H_{01}. The P-value for male test subjects alone across all three treatments is 0.6457, which indicates no significant difference in performance between treatment levels. The combination of these statistical analyses strongly supports the conclusion that the information extraction capability of females is more negatively affected by high amounts of extraneous graphics than that of males in the test population. Perhaps the relatively smaller sample sizes for females (about half the size of male samples) causes some distortion in the statistical analysis, but the overall trend appears to support the conclusion that high amounts of extraneous graphics significantly hinders information extraction accuracy for females. Further research in this area may help to confirm these results.

Sensitivity analysis was also conducted to determine if prior training (either formal or informal) in graphics construction or interpretation, effects information extraction accuracy. Table 6, which follows the same format as Table 5, summarizes the data and results of the training sensitivity analysis. The columns with the letter "T" indicate those who have had training, and the columns with the letters "NT" indicate those who have had no training.

Table 6. Sensitivity Analysis: Training vs. Information Extraction Accuracy

	No Extraneous Graphics		Medium Extraneous Graphics		High Extraneous Graphics		Between Treatments	
	T	NT	T	NT	T	NT	T	NT
Sample Size	22	9	25	11	19	15	65	36
Total Responses	132	54	150	66	114	90		
Correct Responses	124	46	135	59	102	67		
Incorrect Responses	8	8	15	7	12	23		
Incorrect Responses per person	0.36	0.89	0.60	0.64	0.63	1.53		
P-Value	0.0539		0.8923		0.0048		0.3841	0.0442

In the no extraneous graphics and high extraneous graphics conditions, the P-value indicates statistical significance to at least 95% and 99%, respectively. However, in the medium treatment, a statistically significant difference in performance was not observed. The between treatment P-values indicate that those without prior training are responsible for a significant portion of the effect.

By analyzing the incorrect responses per person, it is appears that those with some form of prior graphics training performed in a predictable manner by steadily increasing the incorrect responses per person as the amount of extraneous graphics increases. Meanwhile, those with no prior graphics training demonstrated less predictable performance by doing best in the medium extraneous graphics condition. Perhaps a conservative amount of extraneous graphics actually enhances information extraction accuracy for those without prior training in graphical applications. However, direct comparison of those without training, in the none and medium extraneous graphics treatments, reveals a P-value of 0.4898 (not shown in Table 6), which is not statistically significant. Therefore, conclusive evidence regarding such an effect is not available from this experiment.

As was the case with female participants, the most significant effects on those with no prior training occurred in the high extraneous graphics treatment. The next logical step would be to analyze the performance of females with no prior training. Unfortunately, the number of test subjects in this category is too small to conduct meaningful statistical analysis. An unsubstantiated logical assumption is that females without prior graphics training are most likely to be significantly affected by high amounts of extraneous graphics. Further investigation with samples containing enough females without graphics training for statistical analysis, is necessary to support this assumption.

The final area identified for sensitivity analysis is the potential effects of colorblindness. As discussed previously in this chapter, however, the number and distribution of colorblind test subjects was too small to conduct meaningful statistical analysis. A total of four subjects claimed to be colorblind, three of which viewed the medium extraneous graphics treatment, one saw the high treatment, and none viewed the treatment without extraneous graphics. Due to the predominant use of color graphics associated with modern presentation software, further research should be conducted that specifically addresses the effects of colorblindness on graphical decision-making.

The test questions designed to evaluate decision accuracy (question 7) and decision confidence (question 8), are based on an integer scale ranging from 1 to 7. This scaled response provides interval data that can be analyzed using the same statistical method as with the ordinal data, but also allows for additional analysis techniques that are impractical for ordinal data.

Figures 20 through 22 show the distributions of responses from question 7, for each of the three conditions (none, medium, and high extraneous graphics). These responses are somewhat normally distributed, but the Wilk-Shapiro test for normality, resulted in W-S statistic values between 0.8 and 0.89. To be considered normally distributed, the K-W test statistics must be greater than the statistic's critical values, which are dependent on sample size, for each of the three treatments . The critical value for the no extraneous graphics treatment is 0.929, 0.935 for the medium treatment, and 0.933 for the high treatment. Therefore, the assumption of normally distributed data is not valid. This condition, combined with a lack of equal variances for the responses from question 8, required the use of nonparametric analysis techniques. As was used for the first six test questions, the Kruskal-Wallis test was employed to analyze the responses from questions 7 and 8.

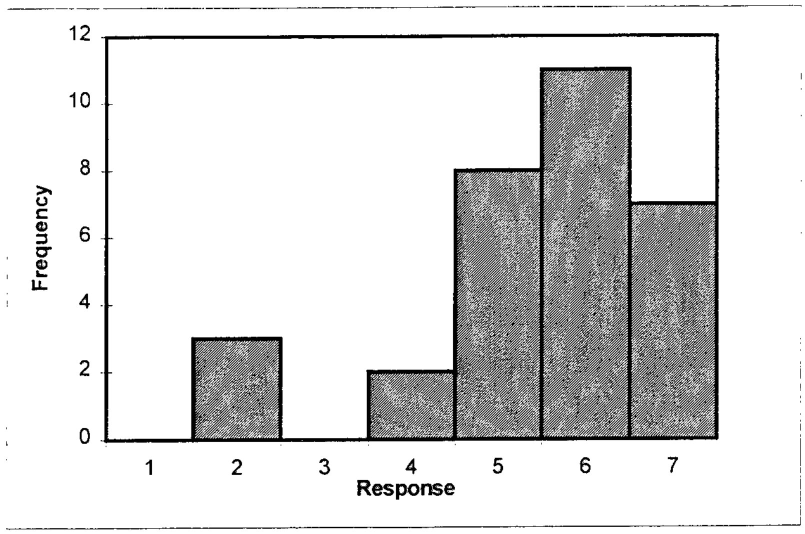


Figure 20. Response Distribution for Decision Accuracy (None)

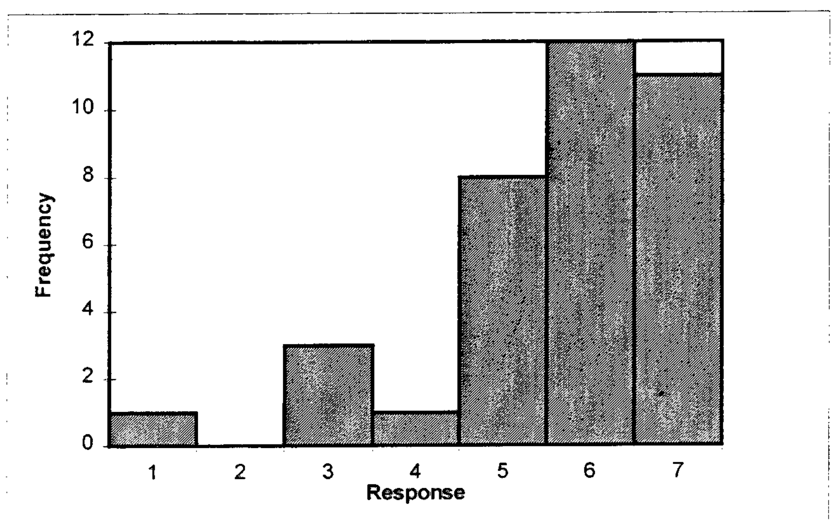


Figure 21. Response Distribution for Decision Accuracy (Medium)

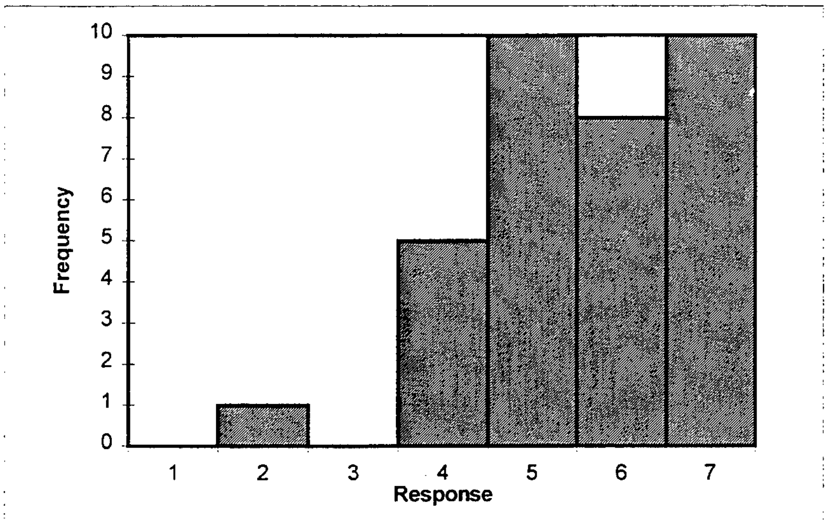


Figure 22. Response Distribution for Decision Accuracy (High)

Table 7 summarizes the descriptive statistics related to the responses obtained for question 7. A 95 percent confidence interval is shown. Although other confidence intervals could have been employed, the 95 percent CI was used as a default value for all calculations.

Table 7. Descriptive Statistics for Decision Accuracy.

Statistic	No Extraneous Graphics	Medium Extraneous Graphics	High Extraneous Graphics
Observations	31	36	34
Sum	169	203	190
Lower 95% CI	4.9257	5.1593	5.1577
Mean	5.4516	5.6389	5.5882
Upper 95% CI	5.9776	6.1184	6.0187
Standard Dev.	1.4338	1.4173	1.2338
Coef. of Variance	26.301	25.134	22.079
Median	6	6	6

The distributions of responses and descriptive statistics are very similar for test question 8, which is used to analyze decision confidence. The complete set of distributions, descriptive statistics, and analysis results are found in Appendices E (decision accuracy) and F (decision confidence). Table 8 provides the resulting Kruskal-Wallis test statistics and P-values for these two response variables. These P-values indicate no statistically significant difference between the three test conditions. Therefore, the null hypothesis for both decision accuracy and decision confidence can not be rejected. This failure to reject the null hypotheses indicates that extraneous graphics have no effect on either of these two response variables.

Table 8. Results for Decision Accuracy and Decision Confidence.

Response Variable	K-W Statistic	P-Value
Decision Accuracy	0.4830	0.7855
Decision Confidence	0.7635	0.6827

A graphical illustration showing the comparison of means and variances (using a 95% confidence interval), is shown in Figure 23. This plot illustrates what was indicated from Table 7 - that the responses across all three levels of extraneous graphics are very similar. The plot shown in Figure 23 corresponds to the responses for decision confidence. The comparison of means plot for decision accuracy is shown in Appendix E. The leftmost bar represents the no extraneous graphics condition and the rightmost bar represents the high extraneous graphics condition.

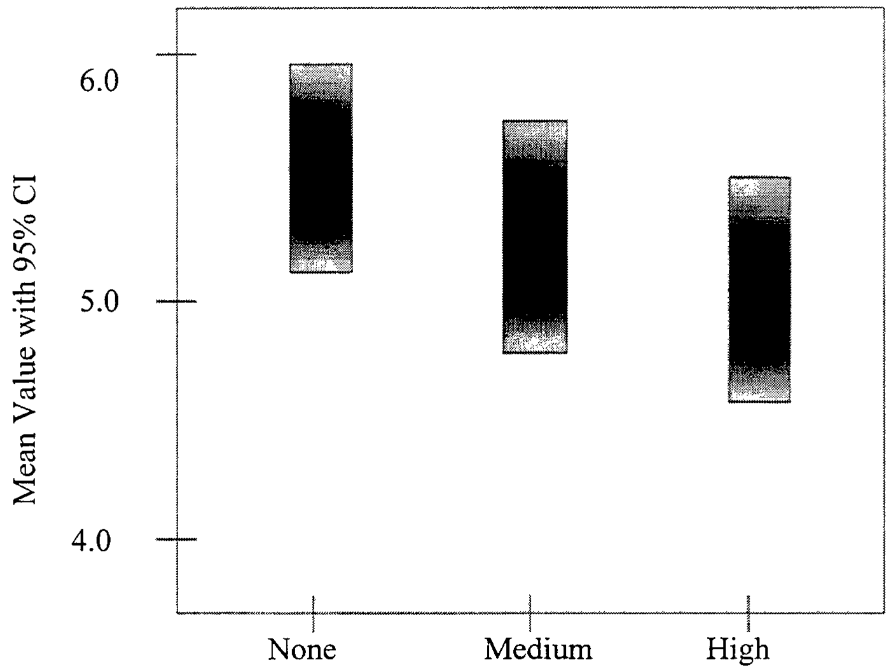


Figure 23. Comparison of Means for Decision Confidence

The mean comparison plot for decision confidence indicates a general trend despite finding no statistical significance in the results. In this case the trends is towards reduced confidence as the level of extraneous graphics increases. Although this trend does not justify an effect, it does seem to indicate the potential for an effect. This trend appears to be positively correlated with information extraction accuracy, meaning that one's confidence in decision-making goes down as the ability to accurately extract information also decreases. Statistical analysis regarding information extraction accuracy showed that test subjects were more adversely affected by extraneous graphics as the amount of these graphics increased. Therefore, despite the lack of significant statistical evidence, the trend analysis indicates a potential relationship between extraction accuracy and decision confidence.

Sensitivity analysis for gender effects on decision accuracy and decision confidence showed a statistically significant effect (up to 96% confidence) in one of the three treatment levels for decision accuracy. Table 9 provides the average responses for males and females in each treatment group, along with the P-value resulting from the Kruskal-Wallis test, which was used to compare the two response distributions. The low P-value (0.0408) for decision accuracy in the high extraneous graphics treatment, indicates that male and female test subjects responded differently.

Table 9. Sensitivity Analysis: Gender vs. Decision Accuracy and Confidence

	No Extraneous Graphics		**Medium Extraneous Graphics**		**High Extraneous Graphics**		**Between Treatments**	
	M	**F**	**M**	**F**	**M**	**F**	**M**	**F**
Sample Size	21	10	25	11	19	15	65	36
Decision Accuracy Average Response	5.571	5.200	5.600	5.727	5.895	5.200		
P-Value	0.4554		0.5797		0.0408		0.588	0.428
Decision Confidence Average Response	5.619	5.400	5.160	5.546	5.316	4.933		
P-Value	0.8527		0.2582		0.3837		0.682	0.441

The average response for males regarding decision accuracy in the high extraneous graphics treatment was 5.895 (on a scale of 1 to 7), and the average response for females was 5.200. This result indicates that females were less likely to "strongly agree" with the decision to use benchmarking after reviewing the presentation (reference Chapter III or Appendices A and B for the decision scenario and decision problem).

This finding corresponds with the results of sensitivity analysis for information extraction accuracy, where females were found to be less accurate in their extraction accuracy in the high extraneous graphics treatment. Therefore, high amounts of extraneous graphics appear to negatively influence information extraction accuracy for females, and subsequently, their decision accuracy is also affected. It is interesting to note, however, that decision confidence among female test subjects was not subsequently affected by the reduced extraction accuracy and decision accuracy. The average response by females, regarding decision confidence, was slightly lower than that of males (4.933 compared to 5.316), but the difference was not statistically significant. Apparently once females make a decision, regardless of how they arrived at it, they have confidence in their decision. Whether or not males demonstrate this same characteristic could not be determined from this experiment since their responses were not significantly affected by extraneous graphics.

Sensitivity analysis for prior graphics training showed similar results. Like the females, the decision accuracy of those test subjects without prior training was slightly affected by the high amounts of extraneous graphics. The P-value comparing the decision accuracy of these two groups (shown in Table 10) indicates a statistical significance for α less than 0.091. Although some may consider this confidence too low for statistical significance, these results seem to correspond to the sensitivity analysis associated with information extraction accuracy. As shown in Table 6, information extraction accuracy also was found to be more adversely affected by high amounts of extraneous graphics for those without prior training than for those with prior training.

These results also seem to parallel the sensitively analysis for females as well (reference discussion in the previous paragraph), indicating that females without graphics training are probably the most likely to be adversely affected by high amounts of extraneous graphics. Again, however, the sample size of females without graphics training is too small for meaningful statistical analysis.

Table 10. Sensitivity Analysis: Training vs. Decision Accuracy and Confidence

	No Extraneous Graphics		Medium Extraneous Graphics		High Extraneous Graphics		Between Treatments	
	T	**NT**	**T**	**NT**	**T**	**NT**	**T**	**NT**
Sample Size	22	9	25	11	19	15	66	35
Decision Accuracy Average Response	5.682	4.889	5.520	5.909	5.842	5.267		
P-Value	0.2984		0.8863		0.0849		0.784	0.231
Decision Confidence Average Response	5.591	5.444	5.040	5.818	5.263	5.000		
P-Value	0.7022		0.2979		0.2782		0.744	0.082

The final area for analysis relates to preferences regarding the use of extraneous graphics. To address this issue, three comparison slides were presented to the test subjects (as shown in Figure 18, and in Appendix A). The subjects were asked to determine which of two slides they felt had a more "professional" appearance. Each of the three possible combinations of treatments was addressed. Question 12 compared example slides from the no extraneous graphics and the medium extraneous graphics conditions. Question 13 compared the medium and high conditions, and question 14 compared the none and high conditions. Table 11 shows the numeric results of this comparison analysis.

Table 11. Results of Preference Comparisons between Treatments

Comparison	None	Medium	High
None vs. Medium	38	61	
Medium vs. High	.	91	8
None vs. High	75		24

The results indicate that the test subjects generally favored a medium amount of extraneous graphics. The medium amount was chosen over both the none and high amounts of extraneous graphics. No extraneous graphics was preferred over a the high amount. Therefore, most participants perceived that extraneous graphics used in moderation (1 - 3 graphics, depending on size, or about 12- 15% of the slides surface area) produces the most professional appearing presentation. However, an overdose of extraneous graphics (more than 3 graphics or more than 20% of the surface area) is worse than none at all.

Summary of Analysis

The primary analysis was conducted in relation to three hypotheses to determine if extraneous graphics affect information extraction accuracy (H_1), decision accuracy (H_2), and decision confidence (H_3). The results of this analysis provided enough statistical evidence to reject H_{01}, and provided the conclusion that extraneous graphics do indeed affect information extraction accuracy. Further analysis showed that extraction accuracy appears to be inversely related to increasing extraneous graphics. The most significant effect was found in the treatment level with the highest amount of extraneous graphics.

Statistical analysis did not provide a basis for rejecting the null hypotheses associated with the other two response variables, and thus, finding that extraneous graphics have no effect on decision accuracy, or decision confidence. However, trend analysis indicated that despite lacking significant statistical evidence of an effect, the experiment demonstrated the potential for an inverse relationship between increasing amounts of extraneous graphics and the two response variables, decision accuracy and decision confidence. The trend analysis also indicated a potential relationship between information extraction accuracy, and the other two response variables. This finding indicates that decreasing information extraction accuracy leads to decreased decision accuracy and confidence.

Sensitivity analysis showed a statistically significant difference in both extraction accuracy and decision accuracy, based on gender and graphics training. The results indicate that females are more readily distracted by large amounts of extraneous graphics, and thus, their decision accuracy was subsequently reduced. It was also shown that high amounts of extraneous graphics are more likely to adversely affect the extraction accuracy and decision accuracy of those test subjects without prior graphics training (either formal or informal). Sensitivity analysis regarding colorblindness was not conducted due to the inhibitively small sample sizes of colorblind test subjects.

The final area of analysis addressed personal preferences regarding the professional appearance of extraneous graphics. The results show that most test subjects felt that a medium amount of graphics looked more professional than either no extraneous graphics or a high amount of extraneous graphics. In the choice between high amounts of extraneous graphics and no extraneous graphics, most preferred none at all.

V. CONCLUSIONS

Extraneous graphics are becoming a standard feature of professional presentations within both the Department of Defense and civilian business community. Understanding how these graphics may influence decision-making is an important research subject. Although little research has been conducted to date regarding the specific influences of extraneous graphics, many previous researchers have studied varies aspects of presentation graphics, particularly the differing effectiveness of various presentation formats such as tables, bar graphs, and pie charts. Opinions differ as to whether graphics enhance or hinder the decision-making process. Some purport the beneficial effects graphics may have on comprehension, interpretation, and aesthetic value, while others consider graphics as a distraction and waste of valuable presentation space. The cumulative results of previous research have been predominately inconclusive in solving this issue.

The potential influence of extraneous graphics on decision-making has been the central theme of this research effort. The primary purpose of this study is to determine if extraneous graphics help or hinder the decision-making process. In pursuit of this topic, three hypothesis were tested to determine if extraneous graphics influence information extraction accuracy, decision accuracy, and decision confidence.

An experiment was conducted to test the proposed hypotheses, and a decision-style presentation was employed as the experimental apparatus. The study was designed to obtain response data pertaining to three response variables: information extraction accuracy, decision accuracy, and decision confidence. Statistical analysis was then conducted to analyze the data and interpret the results.

Summary of Results

Research Question 1. Do extraneous graphics affect information extraction accuracy?

Significant statistical evidence was found to reject the null hypothesis regarding information extraction accuracy. This rejection of the null hypothesis leads to the conclusion that extraneous graphics affect information extraction accuracy. Further analysis indicated that as the amount of extraneous graphics increased the test subject's information extraction accuracy decreased.

Research Question 2. Do extraneous graphics affect decision accuracy?

Research Question 3. Do extraneous graphics affect decision confidence?

The results of the statistical analysis precluded rejecting the null hypothesis for these research questions, and thus, finding that extraneous graphics have no effect on decision accuracy or decision confidence. Preliminary assumptions had theorized that extraction accuracy would be related to decision accuracy and decision confidence. Despite the lack of significant statistical evidence of an effect, trend analysis indicated a potential relationship between decreasing information extraction accuracy, and decreasing decision confidence and decision accuracy. The statistical evidence from this experiment is not significant enough to validate this potential relationship, but it does indicate that further research in this area may be necessary to fully understand the effects of extraneous graphics.

Investigative Question 4. Does gender affect H_1, H_2, or H_3?

Sensitivity analysis showed that high amounts of extraneous graphics appear to negatively influence information extraction accuracy for females (H_1), and subsequently, their decision accuracy (H_2) is also reduced. It is interesting to note, however, that decision confidence (H_3) among female test subjects was not subsequently affected by the reduced extraction accuracy and decision accuracy. Apparently once females make a decision, regardless of how it was arrived at, they have confidence in their decision. Whether or not males demonstrate this same characteristic could not be determined from this experiment since their responses were not significantly affected by extraneous graphics.

Investigative Question 5. Does graphics training affect H_1, H_2, or H_3?

It was shown that high amounts of extraneous graphics are more likely to adversely affect the extraction accuracy and decision accuracy of those test subjects without prior graphics training (either formal or informal). These results seem to parallel the sensitively analysis for females, indicating that females without graphics training are probably the most likely to be adversely affected by high amounts of extraneous graphics. However, the sample size of females without graphics training was too small for meaningful statistical analysis.

Investigative Question 6. Does colorblindness affect H_1, H_2, or H_3?

Sensitivity analysis regarding colorblindness was not conducted due to the inhibitively small sample sizes of colorblind test subjects.

Additional analysis addressed personal preferences regarding the professional appearance of extraneous graphics. The results show that most test subjects felt that a medium amount of graphics looked more professional than either no extraneous graphics or a high amount of extraneous graphics. However, one must be cautious not to dilute presentations with extraneous graphics, since it was found that most people would prefer none at all rather than large doses of extraneous graphics.

Impact of the DoD

Many critical decisions that commit millions of dollars or possibly even human lives, are based on briefings or presentations received every day in the DoD. Ensuring a complete understanding of the potential effects of extraneous graphics on these important decisions is absolutely essential. Within the hierarchical nature of the DoD decision-making structure, decisions made by lower level managers or commanders impact those at higher levels. Presentations used to aid in the decision-making process are used at each successive layer of the hierarchy. Many times the person receiving a briefing at one level, will become the briefer at the next level. Therefore, the information presented to

DoD decision-makers at all levels must be free of influential factors that may adversely affect the decision-making process.

The results of this study indicate that extraneous graphics may significantly hinder the decision-maker's ability to accurately extract information from presentations, especially for female decision-makers. It was found that females are more likely to be distracted by excessive amounts of extraneous graphics. These results are important to decision-makers at all levels in the DoD hierarchy. It applies to both those who employ extraneous graphics in their presentations and to those who receive presentations containing extraneous graphics. Although the results did not indicate a statistically significant effect for decision accuracy and decision confidence for the population as a whole, DoD briefers and briefees should be cautioned not to overuse extraneous graphics. An significant adverse relationship between excessive use of extraneous graphics and extraction accuracy was identified, and a potential adverse relationship between excessive extraneous graphics and decision-making was also identified, but not proven to be statistically significant within the confines of this experiment. While the results suggest generalizability across the DoD population, more conclusive research needs to be accomplished to confirm these results.

Additionally, it was found that presentations with excessive extraneous graphics were judged to be the least professional in appearance. Those without any extraneous graphics were considered to look more professional than those with too much. Therefore, DoD briefers who want to present a professional appearance should confine themselves to moderate usage of extraneous graphics.

Further Research Recommendations

Several areas for further study have become apparent during the course of this research effort, many of which have already been highlighted in previous sections. The primary recommendation is for similar experiments that employ different decision problems. Obviously, the decision problem used in this study is not representative of every possible decision problem within the DoD. The type of decision (operational, administrative, financial, life and death, etc.) may have a dramatic influence on the effects of extraneous graphics.

Another area discussed within this study, is the potential effect that males and females may react differently to extraneous graphics. Further research employing larger samples of female test subjects with other decision problems is needed to verify the results of this experiment. Additionally, the potentially compounded effects of females without graphics training also warrants further investigation.

One area identified for sensitivity analysis in this study was the potential affects of colorblindness on graphics interpretation and decision-making. However, a very small percentage of test subjects were known to be colorblind, and the resultant sample size was far too small for meaningful statistical analysis. A study specifically addressing the potential effects of colorblindness on graphical applications is certainly warranted.

Additional research could also be accomplished on the subject of professional appearance and personal preferences. The analysis conducted within this research effort only scratched the surface of this area. A study dedicated to this issue may provide valuable insights, especially for persuasion-style presentations.

APPENDIX A. DECISION PROBLEM SCENARIO AND INSTRUCTIONS

This Appendix contains the decision problem scenario, instructions for completing the experimental item, an answer sheet, and the demographics questionnaire. Additional instruction also appeared in the experiment presentation, which was viewed on a PC Monitor (a copy of the presentation is provided in Appendix B). The instructions and questionnaire shown in this Appendix were provided as a hand-out on standard bond paper.

The format and font size of the text have been modified to fit within the margin constraints for this thesis document.

Decision Problem

You are the Chief Executive Officer (CEO) of a large industrial firm that produces and distributes ball bearings. Over the past six quarters (1 ½ years) it seems that something is wrong with the firm's distribution process. No one within the company can exactly identify the problem or problems. One possible method of identifying the problem(s), is to benchmark your company against one of the world's best distributors. Benchmarking is the process of comparing your company with a better company (preferably the "world's best") to evaluate your company's performance and identify problem areas.

Based solely on the information contained in the following graphs, you must decide whether to continue working the issue in-house or consult the world's most efficient distributor for assistance in evaluating (i.e. benchmarking) your company's distribution process. Although this benchmark company does not deal in ball bearings, it should be able to identify and help correct the problem.

Currently, your distribution difficulties are suspected to be impacting the company's performance and competitiveness. Your closest competitor in the ball bearing business may be gaining ground on your firm. Although the benchmark in distribution operates in an entirely different business, it is reasonable to assume that this company can help identify your company's distribution problems. The benchmark company, if consulted, is willing to evaluate your processes and aid in correcting your problems - for a fee, of course. Assume that you consider the fee to be reasonable, if your company really needs the help. The decision you are faced with is: do you ask this company to benchmark your process?

Decision Rules

Assume that the six slides in the following presentation contain all the information necessary to make an informed decision. *The computer will automatically advance the slides for you.* You will be given approximately 15 seconds to view each slide. Following the six informational slides, there will be a few additional slides that will ask questions pertaining to the content of the presentation, your final decision regarding the benchmarking issue, and your opinion regarding the quality of the presentation.

To answer the questions at the end of the presentation, please circle your answer on the provided answer sheet.

After finishing the presentation, please complete the short demographics questionnaire on the last page of the answer sheet. Please DO NOT complete the questionnaire until <u>after</u> completing the presentation.

This is not a test and your name will not be requested or recorded.

Thank you in advance for your conscientious participation.

Answer Sheet

Please circle your answer to the corresponding question on each slide

1.	Disagree	Agree
2.	Disagree	Agree
3.	Disagree	Agree
4.	Disagree	Agree
5.	Disagree	Agree
6.	Disagree	Agree

7.	1	2	3	4	5	6	7
8.	1	2	3	4	5	6	7
9.	1	2	3	4	5	6	7
10.	1	2	3	4	5	6	7
11.	1	2	3	4	5	6	7

12.	A	B	
13.	A	B	
14.	A	B	
15.	A	B	C

Questionnaire

This questionnaire contains statements and questions concerning your background information for demographics and your level of experience with graphical information.

1. What is your gender? Female Male

2. What is your education level?

High School	Baccalaureate Degree
Some College	Some Graduate Courses
Associates Degree	Masters Degree
	Doctoral Degree

3. Are you color blind? No Yes
   If Yes, to what degree or level (if known):

4. Have you ever had any training with graph construction or interpretation?

   Circle all that apply
   a) Yes, formal training on graph construction
   b) Yes, formal training on graph interpretation
   c) Yes, informal training on graph construction
   d) Yes, informal training on graph interpretation
   e) No formal or informal training on graph construction or interpretation

5. Have you ever had any training with presentation (or briefing) construction or interpretation?

   Circle all that apply
   a) Yes, formal training on presentation construction
   b) Yes, formal training on presentation interpretation
   c) Yes, informal training on presentation construction
   d) Yes, informal training on presentation interpretation
   e) No formal or informal training on presentation construction or interpretation

5. How often do you use graphical information in decision-making?

Most of the time Frequently Sometimes Rarely Never

6. Were the instructions clear and easy to follow? No Yes
   Comments:

7. The individual experimental questions (those imbedded in the presentation) were easy to understand.

   Strongly Disagree Disagree Neutral Agree Strongly Agree

8. Did you have any previous knowledge of this experiment? No Yes

APPENDIX B. DECISION PROBLEM PRESENTATION

This Appendix contains copies of the three versions of the presentation used in this experiment. The first version is the "no extraneous graphics" treatment, the second version is the "medium extraneous graphics" treatment, and the third version is the "high extraneous graphics" treatment. All three versions contain 22 slides, the first 7 of which are unique to each version. The remaining 15 slides contain the experiment questions and additional instructions. The last 15 slides are identical for all three versions, and therefore, will only be shown once - after the high extraneous graphics treatment.

The presentation employed full color-graphics, so the gray-scale printed version shown here, which has been reduced to fit on the printed page, may be difficult to discern in some instances. Color bar charts were used to convey the important decision-related information in the presentation. These bar charts are color-coded with a key provided on each slide. This color-coding may be difficult to distinguish on the printed copy.

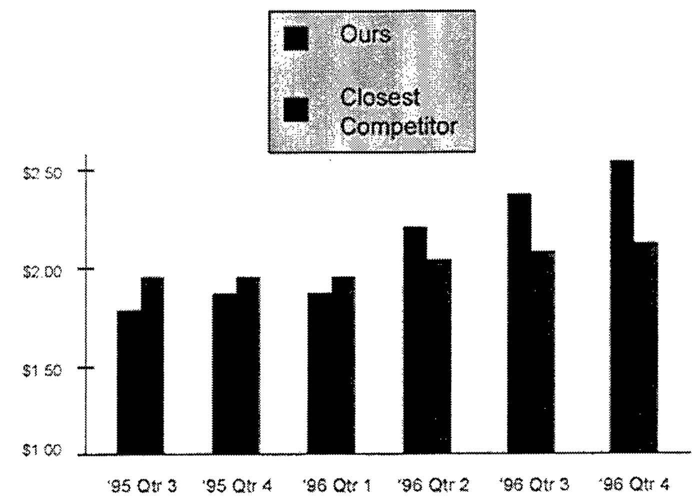
No Extraneous Graphics Version

(First 7 Slides Only)

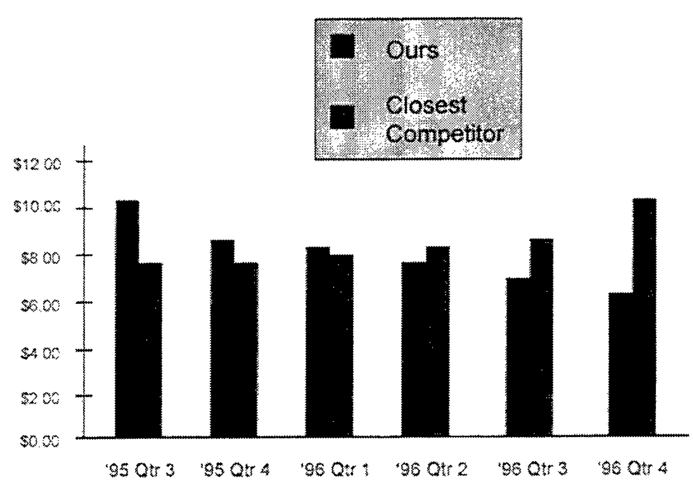
Benchmarking Decision

After reading the Decision Problem and Decision Rules,
press ENTER to continue.

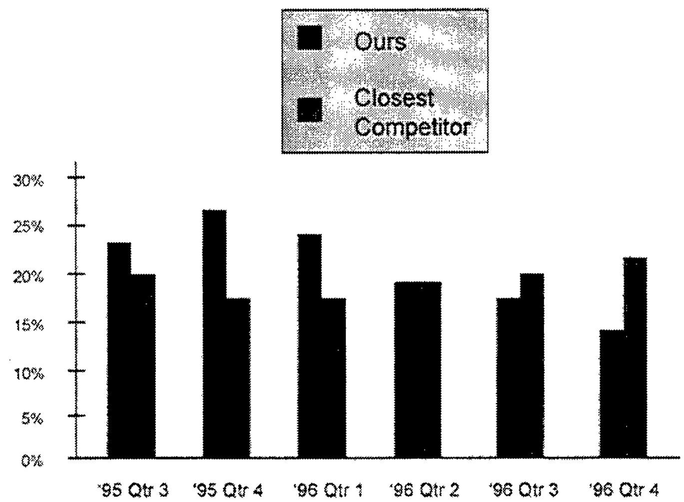
Distribution Cost Per Item



Stock Price Per Share

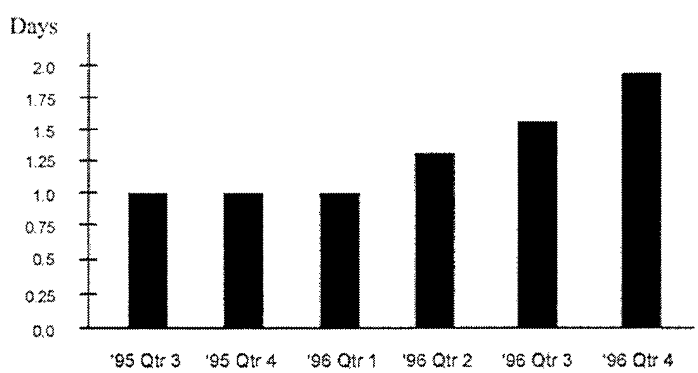


Market Share

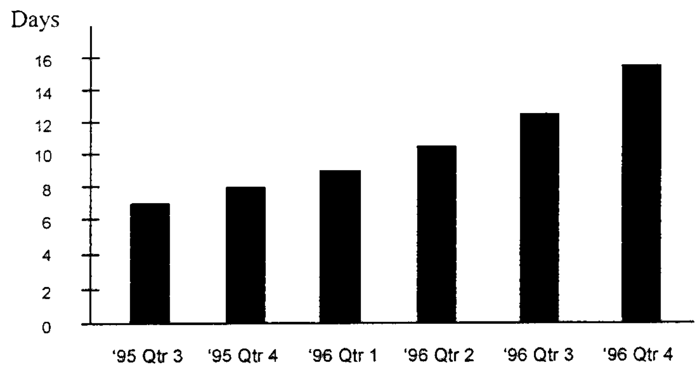


Cycle Time

Time from receipt of invoice to shipment

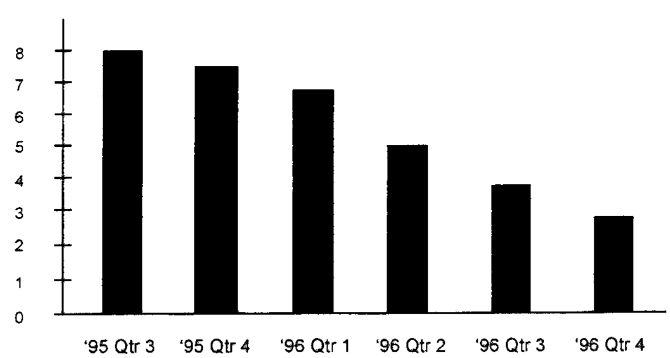


Inventory Turnover Time



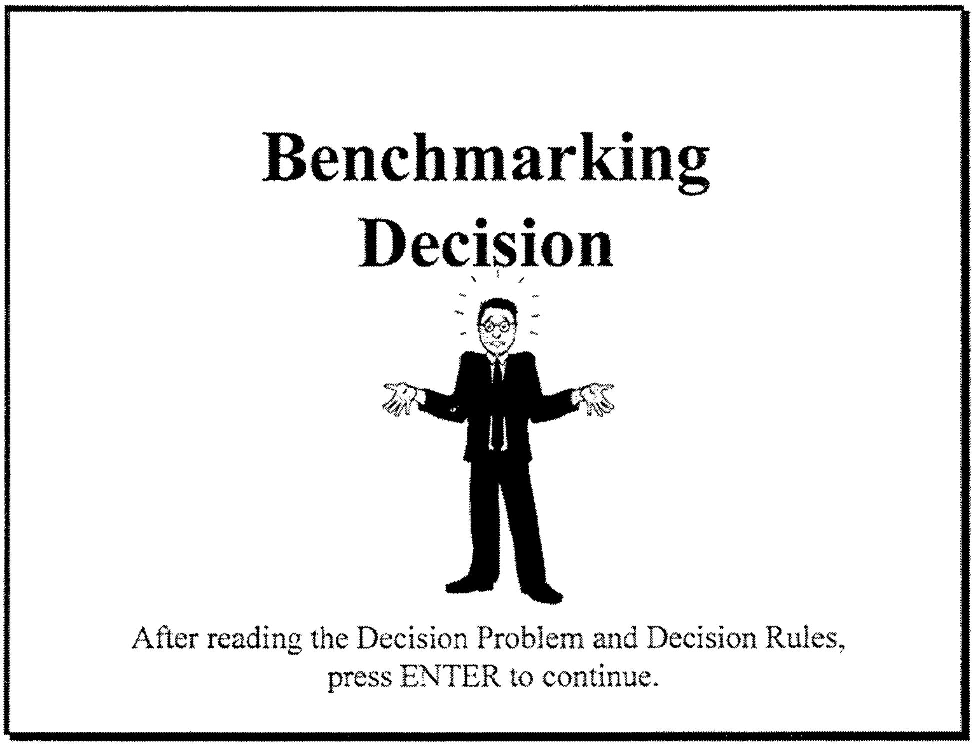
Customer Satisfaction with Distribution/Delivery Time

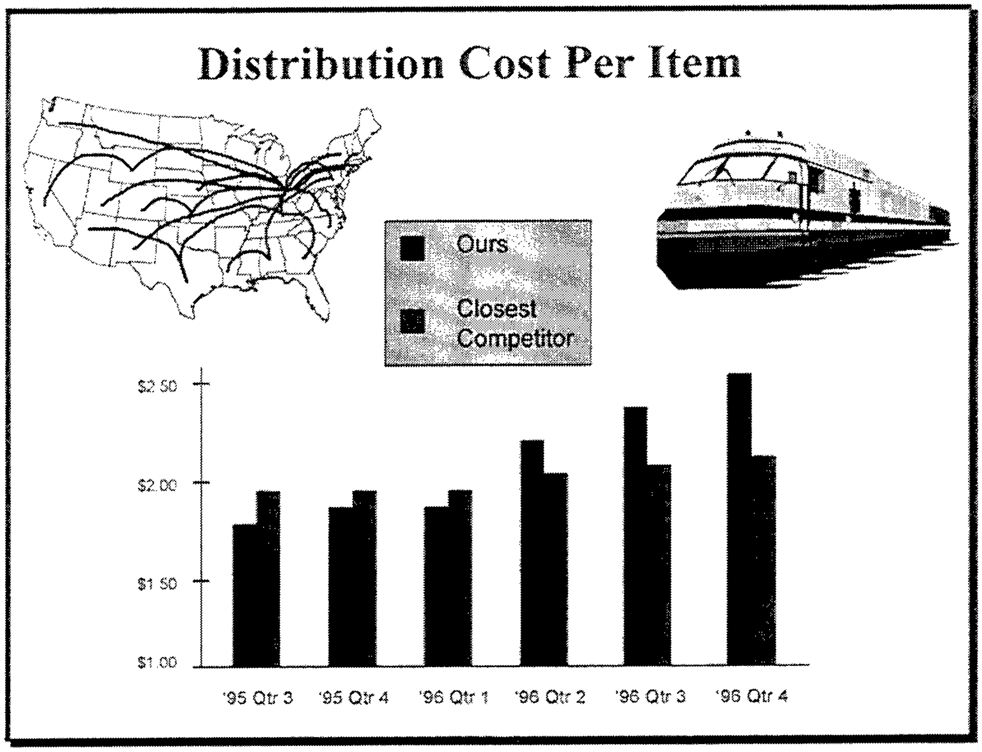
1 = Poor 10 = Excellent

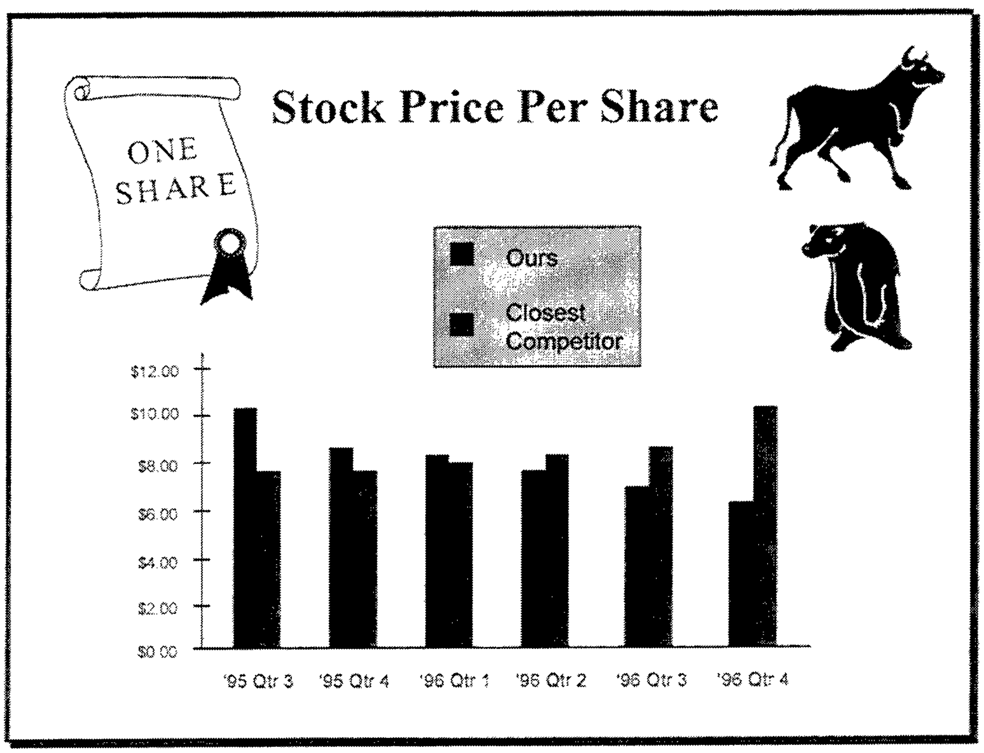


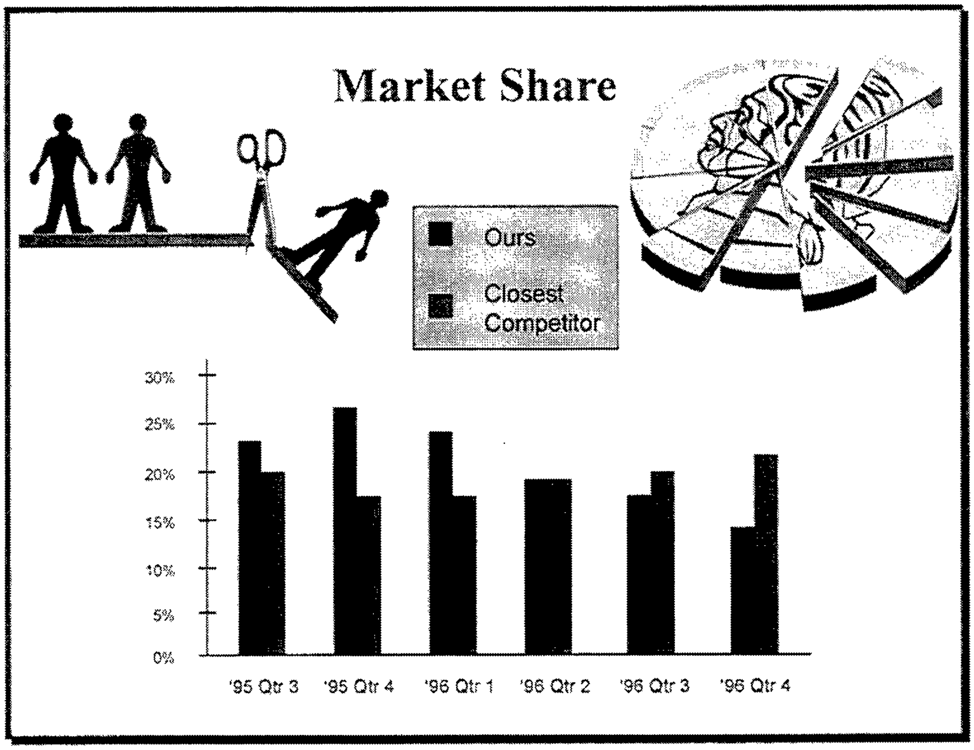
Medium Extraneous Graphics Version

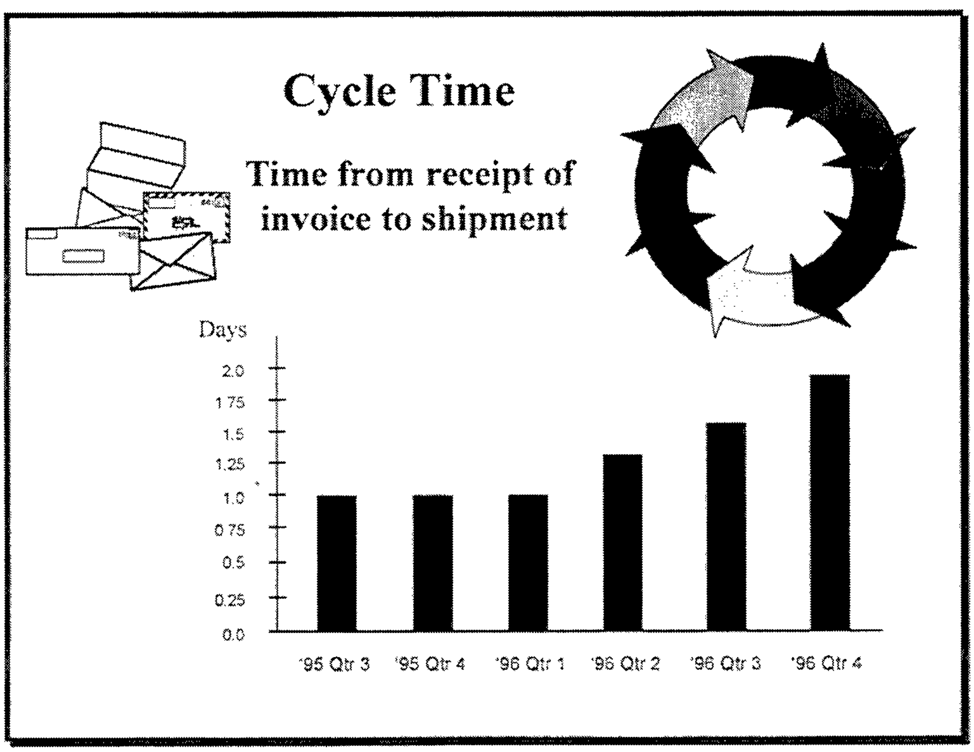
(First 7 Slides Only)



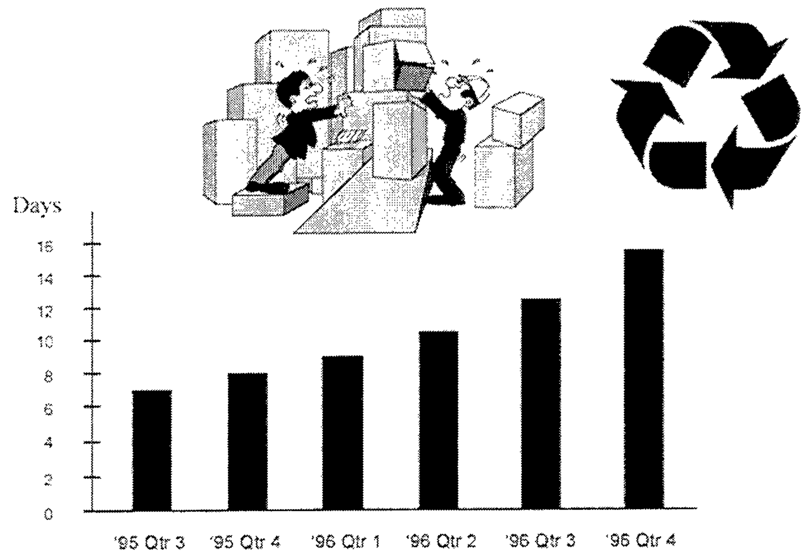

Distribution Cost Per Item
Ours
Closest Competitor
$2.50
$2.00
$1.50
$1.00
'95 Qtr 3
'95 Qtr 4
'96 Qtr 1
'96 Qtr 2
'96 Qtr 3
'96 Qtr 4


Stock Price Per Share
ONE SHARE
Ours
Closest Competitor
$12.00
$10.00
$8.00
$6.00
$4.00
$2.00
$0.00
'95 Qtr 3
'95 Qtr 4
'96 Qtr 1
'96 Qtr 2
'96 Qtr 3
'96 Qtr 4

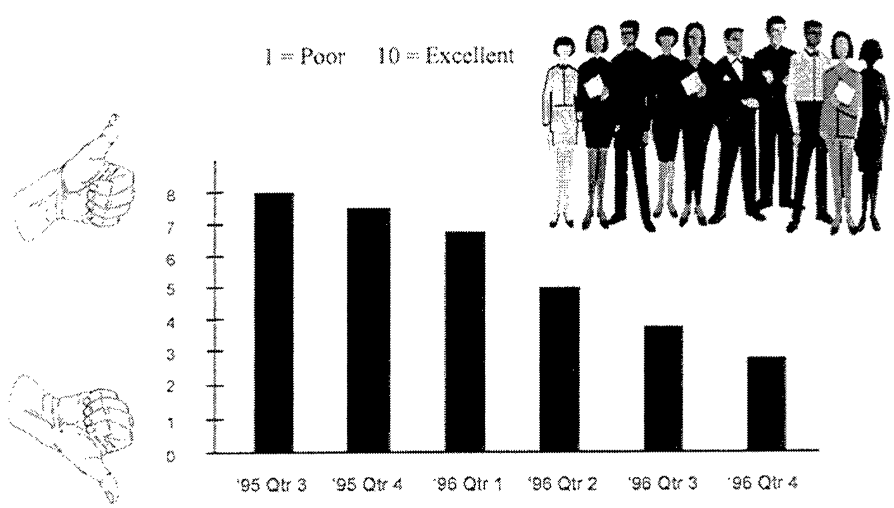

Market Share
Ours
Closest Competitor
30%
25%
20%
15%
10%
5%
0%
'95 Qtr 3
'95 Qtr 4
'96 Qtr 1
'96 Qtr 2
'96 Qtr 3
'96 Qtr 4


Cycle Time
Time from receipt of invoice to shipment
Days
2.0
1.75
1.5
1.25
1.0
0.75
0.5
0.25
0.0
'95 Qtr 3
'95 Qtr 4
'96 Qtr 1
'96 Qtr 2
'96 Qtr 3
'96 Qtr 4

Inventory Turnover Time



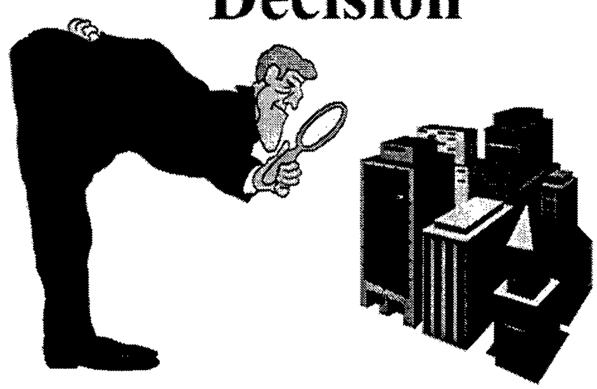
Customer Satisfaction with Distribution/Delivery Time



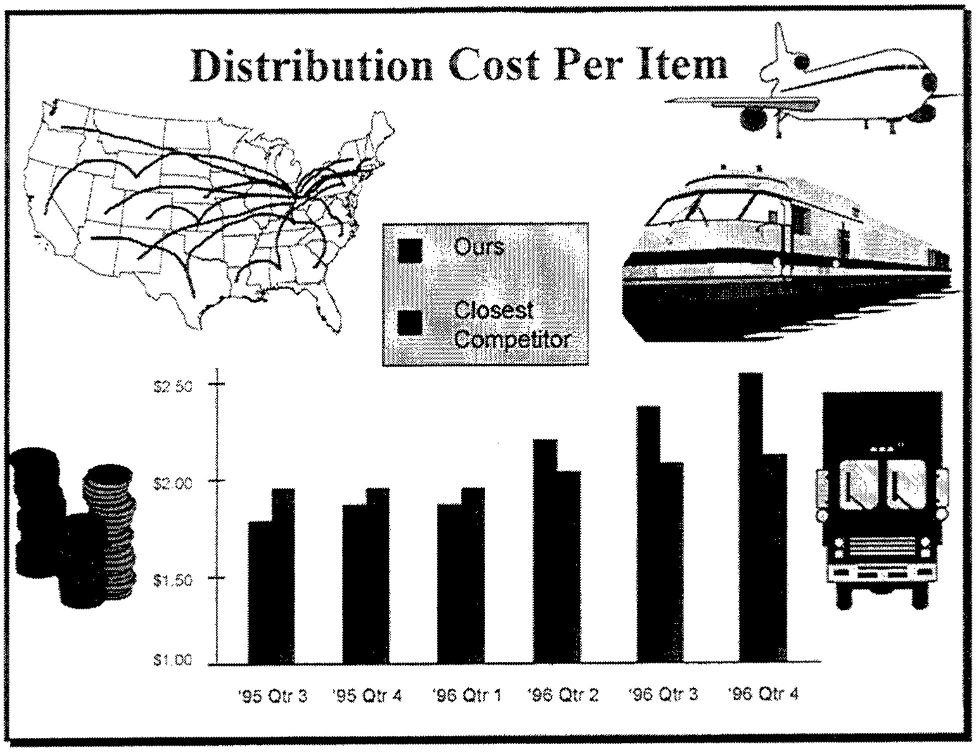
High Extraneous Graphics Version

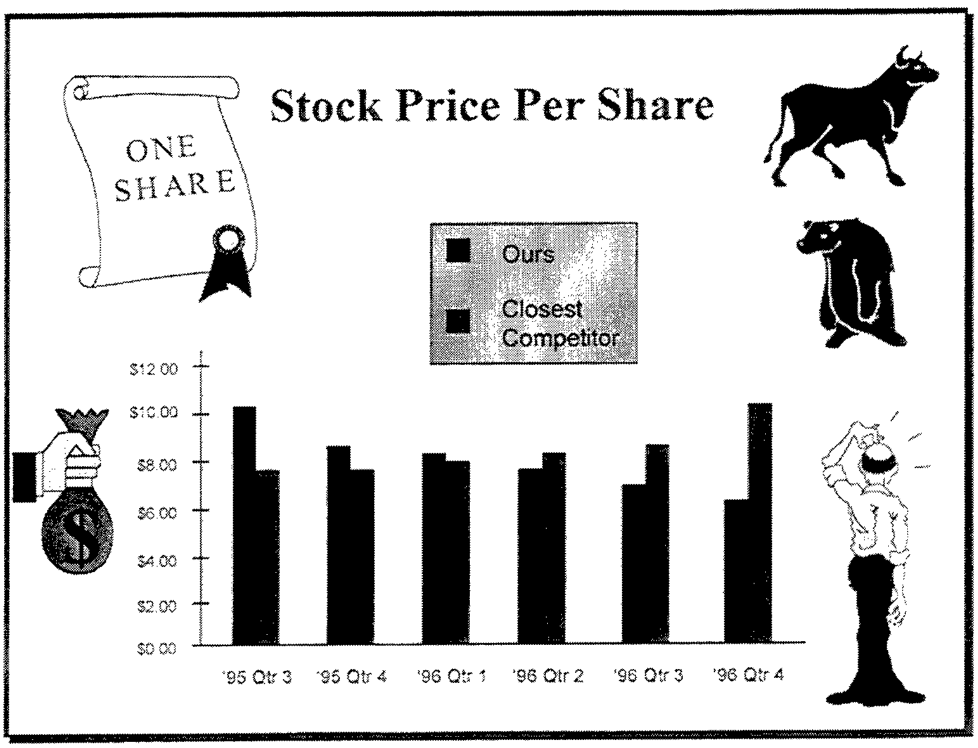
(Entire Presention - Including Question Slides)

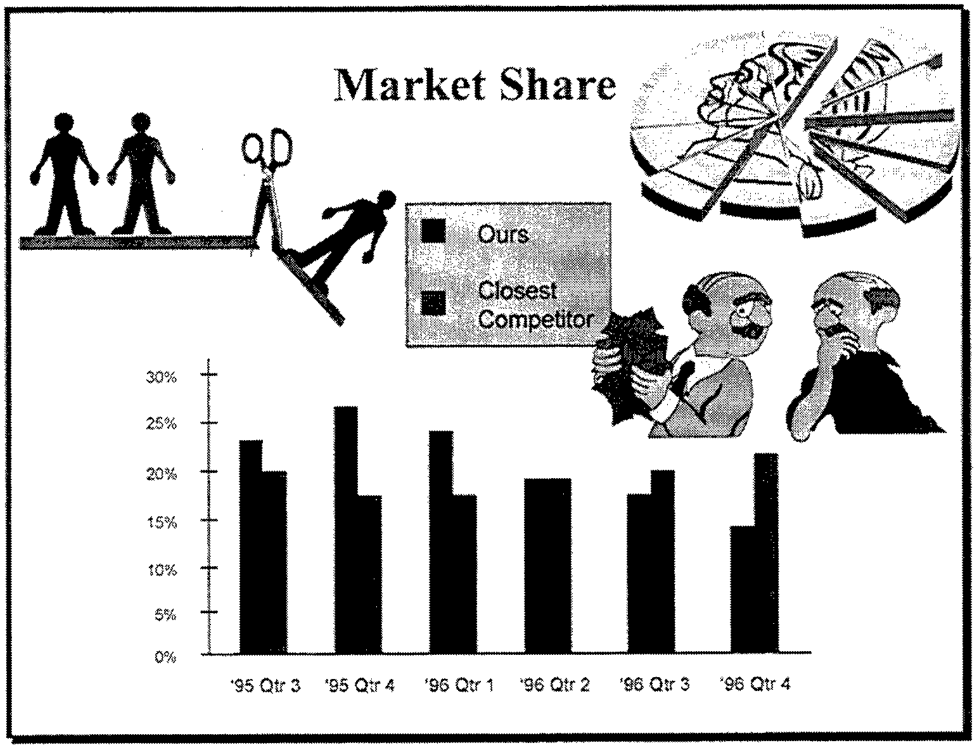
Benchmarking Decision

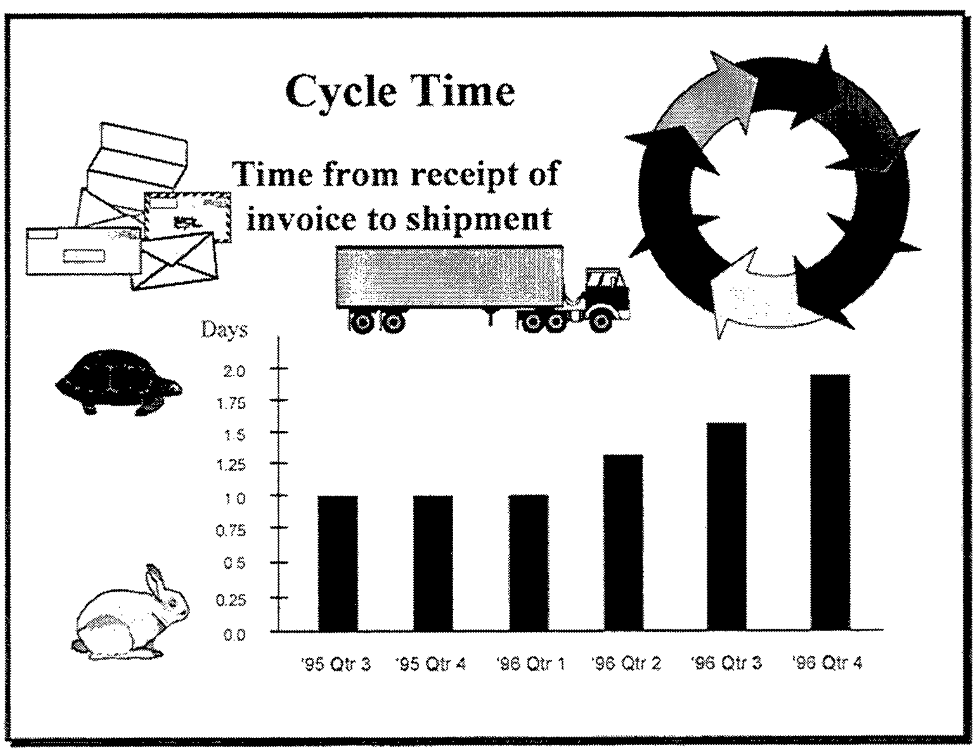


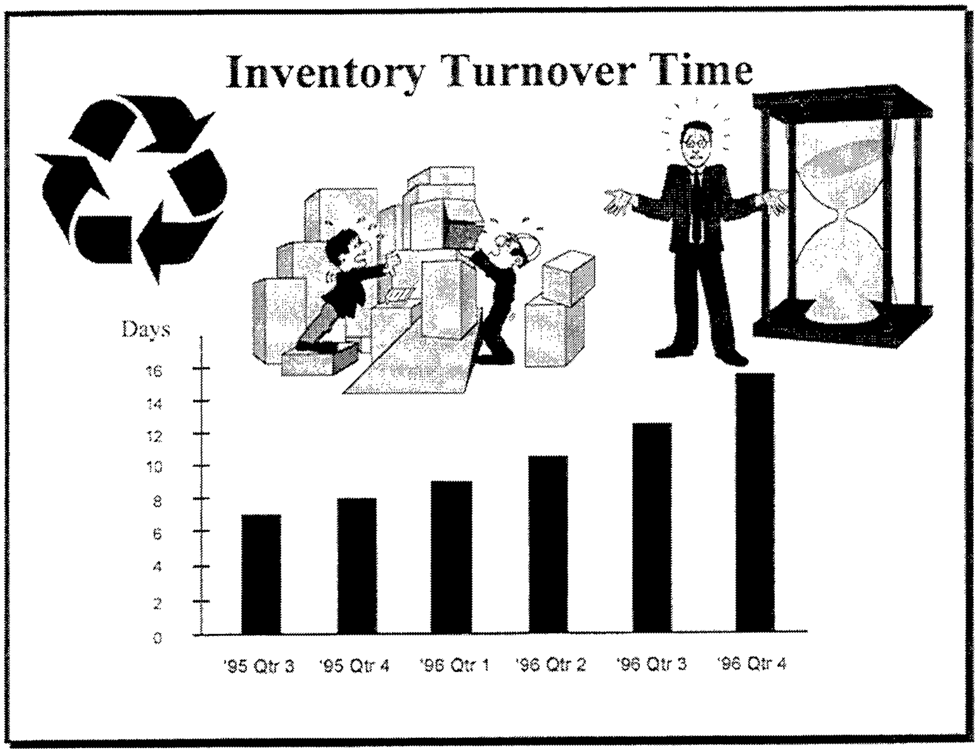
After reading the Decision Problem and Decision Rules, press ENTER to continue.

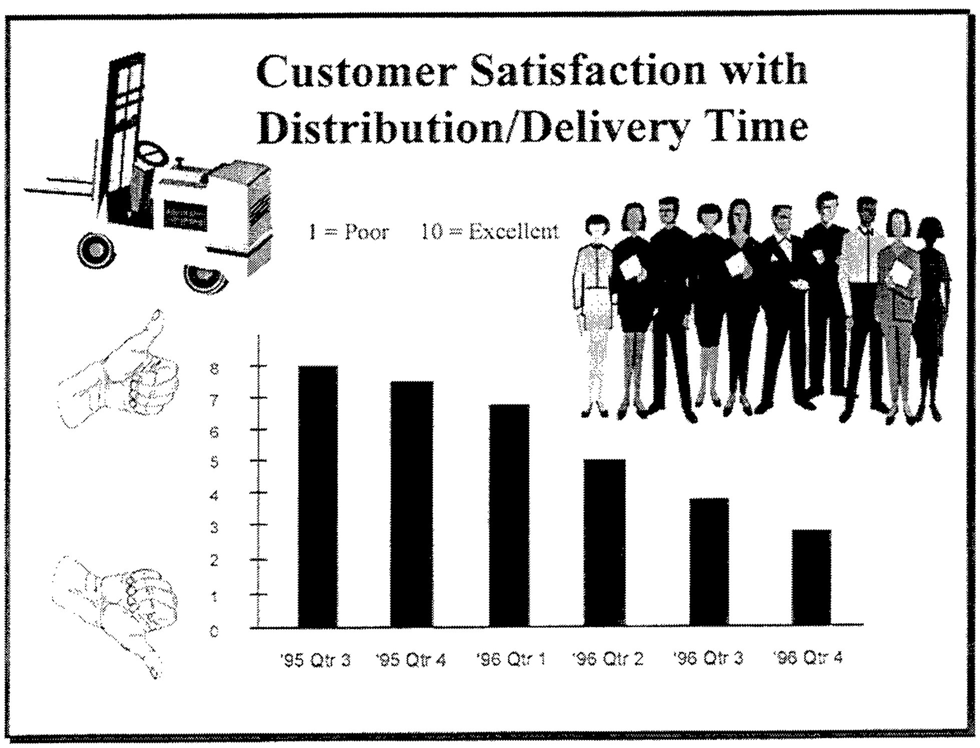

Distribution Cost Per Item
Ours
Closest Competitor
$2.50
$2.00
$1.50
$1.00
'95 Qtr 3
'95 Qtr 4
'96 Qtr 1
'96 Qtr 2
'96 Qtr 3
'96 Qtr 4


Stock Price Per Share
ONE SHARE
Ours
Closest Competitor
$12.00
$10.00
$8.00
$6.00
$4.00
$2.00
$0.00
'95 Qtr 3
'95 Qtr 4
'96 Qtr 1
'96 Qtr 2
'96 Qtr 3
'96 Qtr 4


Market Share
Ours
Closest Competitor
30%
25%
20%
15%
10%
5%
0%
'95 Qtr 3
'95 Qtr 4
'96 Qtr 1
'96 Qtr 2
'96 Qtr 3
'96 Qtr 4


Cycle Time
Time from receipt of invoice to shipment
Days
2.0
1.75
1.5
1.25
1.0
0.75
0.5
0.25
0.0
'95 Qtr 3
'95 Qtr 4
'96 Qtr 1
'96 Qtr 2
'96 Qtr 3
'96 Qtr 4


Inventory Turnover Time
Days
16
14
12
10
8
6
4
2
0
'95 Qtr 3
'95 Qtr 4
'96 Qtr 1
'96 Qtr 2
'96 Qtr 3
'96 Qtr 4


Customer Satisfaction with Distribution/Delivery Time
1 = Poor 10 = Excellent
8
7
6
5
4
3
2
1
0
'95 Qtr 3
'95 Qtr 4
'96 Qtr 1
'96 Qtr 2
'96 Qtr 3
'96 Qtr 4

Please circle your selection on the provided answer sheet

1. Our distribution cost per item has increased more than our closest competitor's distribution cost per item.

Disagree Agree

press ENTER *to continue*

2. Our stock prices have declined over the period under consideration.

Disagree Agree

press ENTER *to continue*

3. Our share of the market has been declining over the past three quarters.

Disagree Agree

press ENTER *to continue*

4. Cycle time has decreased over the past three quarters.

Disagree Agree

press ENTER *to continue*

5. The time needed to go through stocked inventory (turnover time) is declining.

Disagree Agree

press ENTER *to continue*

6. Customer satisfaction with distribution and delivery has declined over the period being considered.

Disagree Agree

press ENTER *to continue*

7. Based solely on the information provided, should the distribution process be evaluated against the world's benchmark distributor?

strongly disagree						strongly agree
1	2	3	4	5	6	7

press ENTER *to continue*

8. Based solely on the information provided, how confident are you that you made the appropriate overall decision?

no confidence						very confident
1	2	3	4	5	6	7

press ENTER *to continue*

9. Do you feel that the presentation contained the appropriate amount of support graphics, such as clip-art, logos, icons, etc., (not including the informational bar charts) for a CEO-level briefing?

strongly disagree						strongly agree
1	2	3	4	5	6	7

10. Was there too much or too little use of support graphics?

too little						too much
1	2	3	4	5	6	7

press ENTER *to continue*

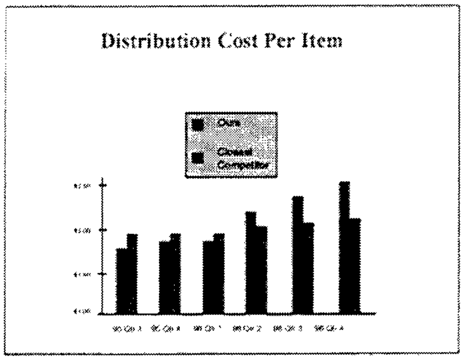
11. What was the impact of the support graphics (clip-art) in the presentation on your ability to understand the information and make a decision?

very detrimental						very helpful
1	2	3	4	5	6	7

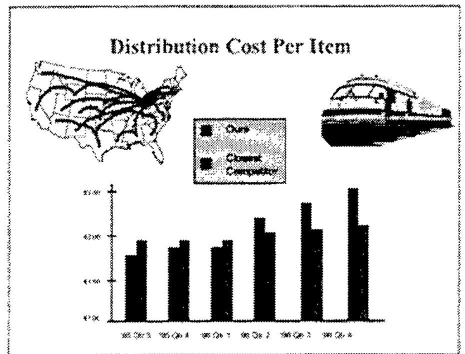
press ENTER *to continue*

12. In your opinion, which of these two slide (A or B) looks more professional?

A.



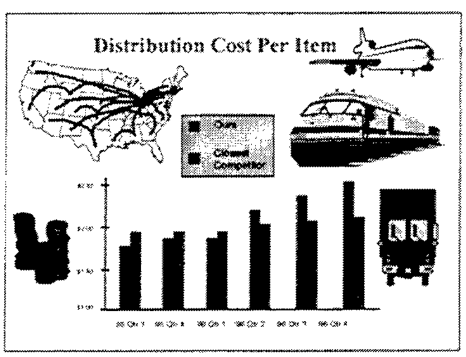
B.



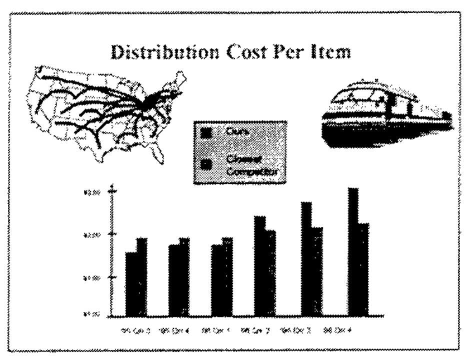
press ENTER *to continue*

13. In your opinion, which of these two slide (A or B) looks more professional?

A.



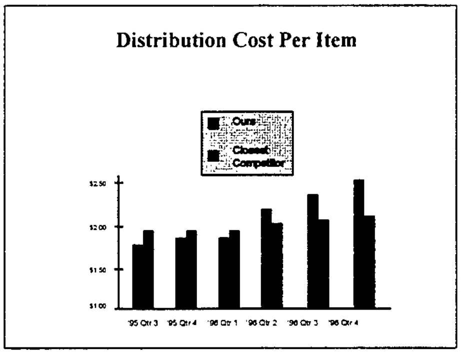
B.



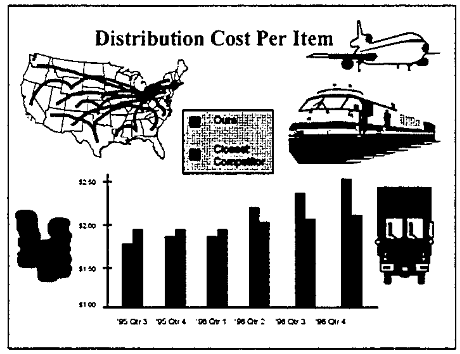
press ENTER *to continue*

14. In your opinion, which of these two slide (A or B) looks more professional?

A.



B.



press ENTER *to continue*

Please circle **C** for number 15 on the answer sheet.

This is to identify which version of the experiment you participated in.

press ENTER *to continue*

This concludes the automated portion of the experiment

Now please complete
the questionnaire

Thank you for your participation

APPENDIX C. RESPONSE DATA

This appendix contains the responses obtained from the test subjects in each of the three treatments (none, medium, and high amounts of extraneous graphics). The data is arranged such that each row represents the set of responses for one test subject. For more details regarding the wording of each question, Appendix B contains the Presentation Questions and Appendix A contains the Demographics Questionnaire. The following key indicates the meaning of the numeric data:

Presentation Questions 1 - 6:	0 = Correct response 1 = Incorrect response
Presentation Questions 7 & 9 (scaled response):	1 = Strongly Disagree 4 = Neutral 7 = Strongly Agree
Presentation Question 8 (scaled response):	1 = No Confidence 4 = Neutral 7 = Very Confident
Presentation Question 10 (scaled response):	1 = Too Little 4 = Neutral 7 = Too Much
Presentation Question 11 (scaled response):	1 = Very Detrimental 4 = Neutral 7 = Very Helpful
Presentation Questions 12 - 14:	0 = Panel A 1 = Panel B
Questionnaire Question 1:	0 = Female 1 = Male

Questionnaire Question 2:	1 = High School 2 = Some College 3 = Associates Degree 4 = Baccalaureate Degree 5 = Some graduate courses 6 = Masters Degree 7 = Doctoral Degree
Questionnaire Question 3:	0 = Not Colorblind 1 = Colorblind
Questionnaire Question 4:	0 = Option "e" (no training) 1 = Option "d" 2 = Option "c" 3 = Option "b" 4 = Option "a"
Questionnaire Question 5:	0 = Option "e" (no training) 1 = Option "d" 2 = Option "c" 3 = Option "b" 4 = Option "a"
Questionnaire Question 6:	1 = Most of the time 2 = Frequently 3 = Sometimes 4 = Rarely 5 = Never
Questionnaire Question 7:	0 = No 1 = Yes
Questionnaire Question 8 (scaled response):	1 = Strongly Disagree 3 = Neutral 5 = Strongly Agree
Questionnaire Question 9:	0 = No 1 = Yes

Response Data for test subjects in the No Extraneous Graphics Treatment

Test	Presentation Questions														Demographics								
Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	6	6	3	3		1	1	0	1	6	0	1, 2	1,2,5	2	1	4	0
2	0	1	0	1	0	1	2	2	1	2		1	1	1	0	3	0	0	0	5	1	4	0
3	0	0	0	0	0	0	6	6	3	3		1	1	1	0	6	0	1,2	1,2	2	1	5	0
4	0	0	0	0	0	0	7	7	4	4		1	1	0	1	5	0	1	1	2	1	5	0
5	0	0	0	0	0	0	6	6	4	4		1	1	0	1	5	0	1,2	1,2	2	1	5	0
6	0	0	0	0	0	0	5	6	6	4		1	1	0	0	5	0	2,3	2	2	1	5	0
7	0	0	0	0	0	0	6	5	3	2		1	1	1	1	4	0	0	2	3	1	4	0
8	0	0	0	0	0	0	7	7	3	3		1	1	1	0	5	0	0	2	3	1	5	0
9	0	0	0	0	0	0	6	6	3	2		1	1	1	1	5	0	2	2	3	1	4	0
10	0	0	0	1	0	0	5	5	4	4		0	1	0	0	3	0	0	0	4	1	3	0
11	0	0	0	0	1	0	5	5	3	3		1	1	0	1	5	0	2	1,2	2	1	4	0
12	0	0	0	0	0	0	7	7	3	3		1	1	1	1	5	0	0	2	2	1	5	0
13	0	0	0	0	0	0	7	6	4	4		0	1	0	0	5	0	0	0	2	1	4	0
14	0	0	0	0	0	0	6	5	5	4		0	1	0	1	5	0	1	1,2	3	0	4	0
15	0	0	0	0	0	0	6	6	3	3		1	1	1	1	5	0	0	1,2	2	1	3	0
16	0	0	0	0	0	0	7	6	4	3		1	1	0	1	5	0	0	0	3	1	4	0
17	0	0	0	0	0	0	6	5	1	1		0	1	0	0	5	0	1,2	1,2	3	1	3	0
18	0	0	0	0	0	0	6	5	5	2		0	1	0	1	5	0	2	2	2	1	4	0
19	0	0	0	1	0	0	2	6	5	5		0	1	0	1	4	0	0	0	2	1	4	0
20	0	0	0	0	0	0	5	6	3	3		0	1	0	1	2	0	0	0	3	1	4	0
21	0	0	0	0	1	0	4	6	5	3		0	1	0	0	4	0	0	0	5	1	4	0
22	0	0	0	1	1	0	5	5	6	6		1	1	0	1	6	0	0	4	4	1	4	0
23	0	0	0	0	0	0	5	5	2	4		0	1	0	1	4	0	1,2	1,2	4	1	4	0
24	0	0	0	0	0	0	7	6	2	2		1	1	1	1	6	0	2	2	2	1	4	0
25	1	0	0	1	0	0	5	3	5	5		0	1	0	1	4	0	2	2	4	1	4	0
26	0	0	0	1	0	0	5	6	6	4		0	1	0	1	5	0	0	0	3	1	4	0
27	0	0	0	1	0	0	7	6	6	3		0	1	0	1	5	0	0	0	4	1	4	0
28	0	0	0	0	0	0	6	5	4	3		1	1	0	1	6	0	1..4	1..4	3	1	4	0
29	0	0	0	1	1	0	4	5	4	6		0	1	0	0	5	0	3	3	3	1	3	0
30	0	0	0	0	0	0	6	6	4	3		0	1	0	0	3	0	1,2	1,2	4	1	4	0
31	0	0	0	1	0	0	2	6	3	2		0	1	0	1	5	0	1	1	4	1	4	0

Response Data for test subjects in the Medium Extraneous Graphics Treatment

Test	Presentation Questions														Demographics								
Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	7	7	7	4	6	1	1	1	0	6	0	1,2	1,2	2	1	5	0
2	1	0	0	0	0	1	7	1	3	2	3	0	1	0	1	5	0	4	4	3	0	1	0
3	0	0	0	0	0	0	5	5	5	3	3	1	1	0	1	6	0	1..4	1..4	2	1	4	0
4	0	0	0	0	0	0	6	6	6	4	6	1	1	1	1	5	0	1,2	1,2	2	1	5	0
5	0	0	0	0	0	0	7	7	6	4	6	1	1	0	0	5	0	1,2	1,2,4	2	1	5	0
6	0	0	0	0	0	0	6	6	4	4	5	1	1	0	1	5	0	0	0	2	1	5	0
7	0	0	0	0	0	0	5	5	4	4	4	1	1	1	1	5	0	0	2	3	1	4	0
8	0	0	0	0	0	0	5	3	5	4	4	1	1	0	1	5	0	3,4	3,4	3	1	4	0
9	0	0	0	0	0	0	6	6	5	4	5	1	1	0	1	5	0	0	0	3	11	3	0
10	0	0	0	0	0	0	6	5	5	4	5	1	1	1	1	6	0	2	4	2	1	4	0
11	0	0	0	0	0	0	7	6	6	4	5	1	1	1	1	5	0	0	2	1	1	5	0
12	0	0	0	0	0	0	5	5	4	4	5	1	1	0	1	4	0	2	2	3	1	4	0
13	0	0	0	0	0	0	6	5	4	4	4	1	1	0	1	3	0	0	2	2	1	4	0
14	0	0	0	0	0	0	6	5	6	5	4	1	1	0	1	4	0	0	0	3	1	4	0
15	0	0	0	0	0	0	6	6	6	4	6	1	1	0	1	4	0	0	2	2	1	4	0
16	0	0	0	0	0	0	7	6	7	4	6	1	0	1	0	3	0	0	0	3	1	4	0
17	0	0	0	0	0	0	6	6	6	4	5	1	1	0	0	2	0	0	0	4	1	3	0
18	0	0	0	0	0	0	7	7	6	4	6	1	1	1	0	4	0	0	0	4	1	4	0
19	0	0	0	0	0	0	6	5	3	5	4	0	1	0	0	5	0	0	0	4	1	4	0
20	0	0	0	0	0	0	5	6	3	3	2	0	1	0	0	2	0	0	0	5	1	3	0
21	1	0	0	1	0	0	7	7	7	5	7	0	1	0	0	6	0	1,2	2,3	4	1	5	0
22	0	0	0	0	0	0	7	7	7	3	6	0	1	0	1	5	1	4	4	2	1	5	0
23	0	0	0	0	0	0	7	6	5	5	3	1	1	0	1	5	0	2	3,4	2	1	4	0
24	0	0	1	0	0	0	3	1	1	4	4	0	1	0	0	5	0	2	2	3	1	2	0
25	0	0	1	1	1	1	5	6	7	6	7	1	0	1	1	4	0	0	0	3	1	4	0
26	0	0	0	0	0	0	7	6	1	7	3	0	1	0	1	5	1	1..4	1..4	2	1	5	0
27	0	0	0	0	0	0	6	6	2	4	5	1	1	1	1	6	0	1,2	1,2	2	1	5	0
28	0	0	0	0	0	0	7	7	4	4	4	1	1	0	1	7	0	0	2,4	3	1	5	0
29	0	1	0	0	1	0	3	5	1	7	1	0	1	0	1	5	0	1,2	1,2	3	1	3	0
30	1	1	0	1	1	0	5	5	6	7	6	1	0	1	0	5	1	1,2	1,2	3	1	5	0
31	0	0	1	1	0	0	6	6	5	6	6	0	1	0	1	4	0	0	0	4	0	2	0
32	0	1	0	0	0	0	5	5	5	5	6	0	1	0	1	3	0	0	0	5	0	2	0
33	0	0	0	0	0	0	4	3	6	5	4	0	1	0	1	6	0	1,2	2,3	2	1	4	0
34	1	0	0	1	0	0	6	2	6	5	4	1	1	0	1	6	0	1,2	0	4	1	4	0
35	0	0	0	1	0	0	3	4	3	4	6	0	1	0	0	2	0	0	1	4	1	4	0
36	0	0	0	1	0	0	1	6	5	4	4	1	1	0	1	4	0	3,4	3,4	3	1	4	0

Response Data for test subjects in the High Extraneous Graphics Treatment

Test	Presentation Questions														Demographics								
Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	7	7	7	4	7	1	1	1	1	5	0	0	0	2	1	5	0
2	0	0	0	1	0	0	4	2	2	6	3	0	1	0	1	5	0	0	0	4	1	3	0
3	0	0	0	1	0	0	5	5	3	5	3	1	1	0	1	3	0	0	0	3	1	4	0
4	0	0	0	0	0	0	7	6	5	4	5	1	1	1	1	5	0	0	0	4	1	4	0
5	1	0	0	0	0	0	6	6	4	4	4	0	1	0	1	5	1	0	0	3	1	4	0
6	0	0	0	0	0	0	7	7	6	4	6	1	0	1	0	5	0	2	2	2	1	5	0
7	0	0	0	0	1	0	6	6	3	5	3	0	1	0	1	2	0	0	0	4	1	4	0
8	0	0	0	0	0	0	7	6	6	4	5	1	0	1	1	4	0	0	2	2	1	5	0
9	0	0	0	1	0	0	5	5	5	5	5	1	0	1	0	4	0	0	0	2	1	5	0
10	0	0	0	0	0	0	7	6	3	6	3	1	1	0	1	4	0	2	2	2	1	5	0
11	0	0	0	0	1	0	6	5	3	5	4	1	1	0	1	5	0	0	2	3	1	4	0
12	0	0	0	0	0	0	6	6	5	5	3	1	1	0	1	4	0	0	2	3	1	4	0
13	1	0	0	0	1	0	4	4	3	6	3	1	1	0	0	3	0	0	0	3	1	4	0
14	0	0	0	0	0	0	7	7	4	6	3	1	1	0	1	5	0	4	2	3	1	4	0
15	0	0	0	0	0	0	7	7	5	4	5	0	1	0	0	2	0	2	0	4	1	5	0
16	1	0	1	0	0	0	6	7	1	7	1	0	1	0	1	4	0	1,2	1,2	4	1	5	0
17	0	0	0	0	0	0	7	5	6	4	6	1	1	1	1	4	0	0	0	3	1	5	0
18	0	0	0	0	0	0	6	6	2	5	5	0	1	0	1	5	0	3,4	3,4	4	1	4	0
19	0	0	0	1	0	0	7	6	5	7	6	1	0	1	1	6	0	3	2	3	1	4	0
20	0	1	0	0	0	1	5	5	5	4	6	1	1	0	0	6	0	0	0	5	1	4	0
21	0	0	1	1	1	1	4	5	5	4	4	0	1	0	0	2	0	0	0	5	1	5	0
22	1	0	0	1	0	0	5	5	3	6	3	1	1	0	0	2	0	0	0	3	1	4	0
23	0	0	0	0	0	0	5	5	1	7	1	0	1	0	1	6	0	1	1,2	2	1	4	0
24	1	0	0	0	0	0	2	3	2	4	5	1	1	0	1	7	0	2	2	2	1	5	0
25	0	0	0	1	1	0	5	4	6	5	7	1	0	0	0	2	0	0	0	3	1	5	0
26	0	0	0	0	0	0	4	2	4	4	5	1	1	0	0	4	0	1	2	4	1	4	0
27	0	0	0	0	0	0	5	4	6	4	5	1	1	0	0	4	0	2,4	2,3,4	3	1	3	0
28	1	1	0	0	1	0	5	4	5	5	3	1	1	0	0	4	0	0	0	4	1	3	0
29	1	0	0	0	0	0	6	5	5	6	3	0	1	0	1	5	0	1,2,3	2,3	3	1	3	0
30	0	0	0	1	1	0	7	7	4	7	4	0	1	0	0	4	0	4	4	5	1	4	0
31	0	0	0	0	0	0	6	7	1	5	3	0	1	0	0	6	0	1,4	2	4	1	2	0
32	0	0	0	1	1	0	5	2	5	5	6	1	1	0	0	1	0	2	2	3	1	4	0
33	0	0	0	1	1	0	5	2	5	5	6	1	1	0	1	4	0	2	2	2	1	4	0
34	0	1	0	1	0	1	4	6	6	5	5	0	1	0	0	2	0	0	0	4	1	4	0

APPENDIX D. STATISTICAL ANALYSIS FOR H_1

This appendix contains the results of the statistical analysis for hypothesis H_1, regarding information extraction accuracy. There were six test question pertaining to information extraction accuracy (reference Chapter III or Appendix B) - one for each of the six informational slides in the presentation. The response data used to conduct the analysis is shown in Appendix C, and the statistical analysis was accomplished using the software program Statistix®.

The response data for these six test questions was first analyzed as a whole (comparing the total number of incorrect responses for all six questions across each treatment level), and then individually for each question. The results of the statistical analysis comparing of totals, are summarized in Table 3 (found in Chapter IV). An additional table, shown in this appendix, summarizes the results when comparing each question individually across the three treatment levels. A discussion of these results is found in Chapter IV.

Because the results from comparing the totals indicated a statistically significant difference, further analysis was conducted to determine which treatment or treatments were different. To do so, each possible pair of treatments (none vs. medium, medium vs. high, and none vs. high) was analyzed.

DESCRIPTIVE STATISTICS

VARIABLE	N	MEAN	SD
NONE1_6	186	0.0860	0.2812
MED1_6	216	0.1019	0.3032
HIGH1_6	204	0.1716	0.3779

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR QUESTIONS 1-6

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE1_6	293.1	186
MED1_6	297.9	216
HIGH1_6	319.0	204
TOTAL	303.5	606

KRUSKAL-WALLIS STATISTIC 7.8045
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.0202

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	76041.5	38020.7	3.94	0.0196
WITHIN	603	5818672	9649.54		
TOTAL	605	5894713			

TOTAL NUMBER OF VALUES THAT WERE TIED 606
MAX. DIFF. ALLOWED BETWEEN TIES 0.00001

CHI-SQUARE TEST FOR HETEROGENEITY OR INDEPENDENCE (QUESTIONS 1-6)

CASE		VARIABLE RIGHT	WRONG	
NONE	OBSERVED	170	16	186
	EXPECTED	163.59	22.41	
	CELL CHI-SQ	0.25	1.83	
MED	OBSERVED	194	22	216
	EXPECTED	189.98	26.02	
	CELL CHI-SQ	0.09	0.62	
HIGH	OBSERVED	169	35	204
	EXPECTED	179.43	24.57	
	CELL CHI-SQ	0.61	4.42	
		533	73	606

OVERALL CHI-SQUARE 7.82
P-VALUE 0.0201
DEGREES OF FREEDOM 2

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR NONE1_6 VS. MED1_6

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE1_6	199.8	186
MED1_6	203.0	216
TOTAL	201.5	402

KRUSKAL-WALLIS STATISTIC	0.2919
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.5890

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	1011.84	1011.84	0.29	0.5896
WITHIN	400	1389104	3472.76		
TOTAL	401	1390116			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR MED1_6 VS. HIGH1_6

VARIABLE	MEAN RANK	SAMPLE SIZE
MED1_6	203.4	216
HIGH1_6	218.0	204
TOTAL	210.5	420

KRUSKAL-WALLIS STATISTIC	4.3370
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.0373

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	22487.8	22487.8	4.37	0.0371
WITHIN	418	2150067	5143.70		
TOTAL	419	2172555			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR NONE1_6 VS. HIGH1_6

VARIABLE	MEAN RANK	SAMPLE SIZE
HIGH1_6	203.5	204
NONE1_6	186.8	186
TOTAL	195.5	390

KRUSKAL-WALLIS STATISTIC	6.2479
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.0124

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	27074.4	27074.4	6.33	0.0122
WITHIN	388	1658603	4274.75		
TOTAL	389	1685678			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR QUESTION 1

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_Q1	46.6	31
MED_Q1	50.6	36
HIGH_Q1	55.4	34
TOTAL	51.0	101

KRUSKAL-WALLIS STATISTIC	4.6541
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.0976

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	1255.07	627.535	2.39	0.0947
WITHIN	98	25711.9	262.367		
TOTAL	100	26967.0			

TOTAL NUMBER OF VALUES THAT WERE TIED 101
MAX. DIFF. ALLOWED BETWEEN TIES 0.00001

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR QUESTION 2

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_Q2	49.1	31
MED_Q2	51.7	36
HIGH_Q2	52.0	34
TOTAL	51.0	101

KRUSKAL-WALLIS STATISTIC	0.9488
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.6222

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	157.645	78.8224	0.47	0.6325
WITHIN	98	16456.9	167.927		
TOTAL	100	16614.5			

TOTAL NUMBER OF VALUES THAT WERE TIED 101
MAX. DIFF. ALLOWED BETWEEN TIES 0.00001

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR QUESTION 3

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_Q3	48.5	31
MED_Q3	52.7	36
HIGH_Q3	51.5	34
TOTAL	51.0	101

KRUSKAL-WALLIS STATISTIC 2.5276
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.2826

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	306.342	153.171	1.27	0.2848
WITHIN	98	11813.7	120.548		
TOTAL	100	12120.0			

TOTAL NUMBER OF VALUES THAT WERE TIED 101
MAX. DIFF. ALLOWED BETWEEN TIES 0.00001

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR QUESTION 4

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_Q4	52.2	31
MED_Q4	47.3	36
HIGH_Q4	53.8	34
TOTAL	51.0	101

KRUSKAL-WALLIS STATISTIC 1.5924
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.4510

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	803.370	401.685	0.79	0.4590
WITHIN	98	49646.1	506.593		
TOTAL	100	50449.5			

TOTAL NUMBER OF VALUES THAT WERE TIED 101
MAX. DIFF. ALLOWED BETWEEN TIES 0.00001

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR QUESTION 5

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_Q5	49.5	31
MED_Q5	47.2	36
HIGH_Q5	56.4	34
TOTAL	51.0	101

KRUSKAL-WALLIS STATISTIC 4.5586
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.1024

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	1565.42	782.708	2.34	0.0995
WITHIN	98	32774.6	334.435		
TOTAL	100	34340.0			

TOTAL NUMBER OF VALUES THAT WERE TIED 101
MAX. DIFF. ALLOWED BETWEEN TIES 0.00001

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR QUESTION 6

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_Q6	49.6	31
MED_Q6	50.8	36
HIGH_Q6	52.5	34
TOTAL	51.0	101

KRUSKAL-WALLIS STATISTIC 0.9150
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.6329

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	131.693	65.8467	0.45	0.6430
WITHIN	98	14260.8	145.518		
TOTAL	100	14392.5			

TOTAL NUMBER OF VALUES THAT WERE TIED 101
MAX. DIFF. ALLOWED BETWEEN TIES 0.00001

Summary of Individual Question Data and Results for Information Extraction Accuracy.

	No Extraneous Graphics	Medium Extraneous Graphics	High Extraneous Graphics	K-W Test Statistic	P-Value
Question 1	1	4	7	4.6541	0.0976
Question 2	1	3	3	0.9488	0.6222
Question 3	0	3	2	2.5276	0.2826
Question 4	9	7	11	1.5924	0.4510
Question 5	4	3	9	4.5586	0.1024
Question 6	1	2	3	0.9150	0.6329
Total	16	22	35	7.8045	0.0202

APPENDIX E. STATISTICAL ANALYSIS FOR H_2

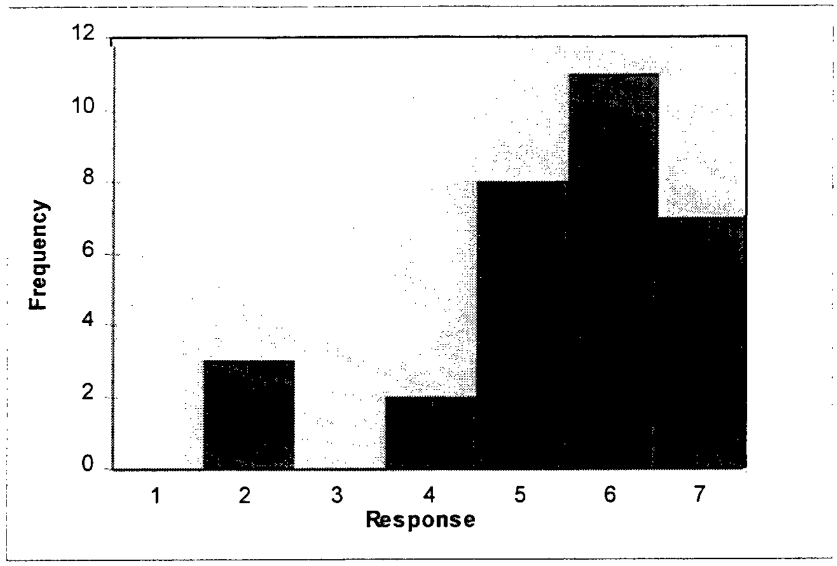
This appendix contains the results of the statistical analysis for hypothesis H_2, regarding decision accuracy. Test question seven pertained to decision accuracy (reference Chapter III or Appendix B). The response data used to conduct the analysis is shown in Appendix C, and the statistical analysis was accomplished using the software program Statistix®.

A table of Descriptive Statistics are provided, followed by histograms illustrating the distribution of responses for each treatment. Finally, the results of the Kruskal-Wallis test are shown. A discussion of these results is found in Chapter IV.

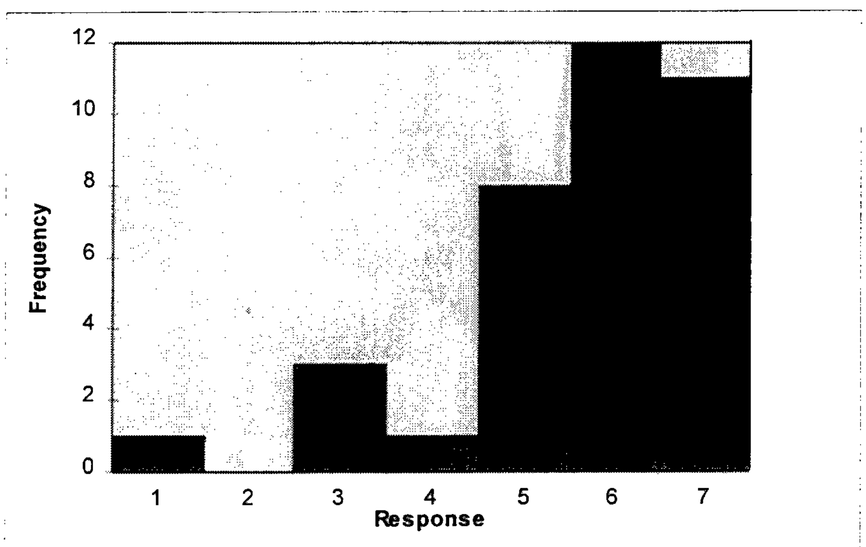
DESCRIPTIVE STATISTICS

	NONE_Q7	MED_Q7	HIGH_Q7
N	31	36	34
SUM	169	203	190
LO 95% CI	4.9257	5.1593	5.1577
MEAN	5.4516	5.6389	5.5882
UP 95% CI	5.9776	6.1184	6.0187
SD	1.4338	1.4173	1.2338
VARIANCE	2.0559	2.0087	1.5223
SE MEAN	0.2575	0.2362	0.2116
C.V.	26.301	25.134	22.079
MEDIAN	6.0000	6.0000	6.0000

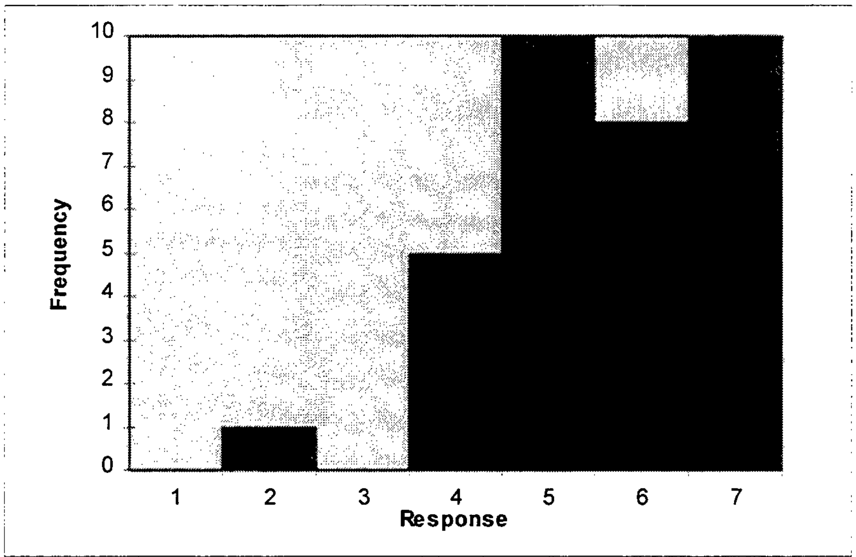
Distribution for Question 7 (No Extraneous Graphics)



Distribution for Question 7 (Medium Extraneous Graphics)



Distribution for Question 7 (High Extraneous Graphics)



KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR: NONE_Q7 MED_Q7 HIGH_Q7

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_Q7	49.1	31
MED_Q7	53.6	36
HIGH_Q7	50.0	34
TOTAL	51.0	101

KRUSKAL-WALLIS STATISTIC 0.4830
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.7855

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	386.540	193.270	0.24	0.7908
WITHIN	98	79645.0	812.704		
TOTAL	100	80031.5			

TOTAL NUMBER OF VALUES THAT WERE TIED 100
MAX. DIFF. ALLOWED BETWEEN TIES 0.00001

APPENDIX F. STATISTICAL ANALYSIS FOR H_3

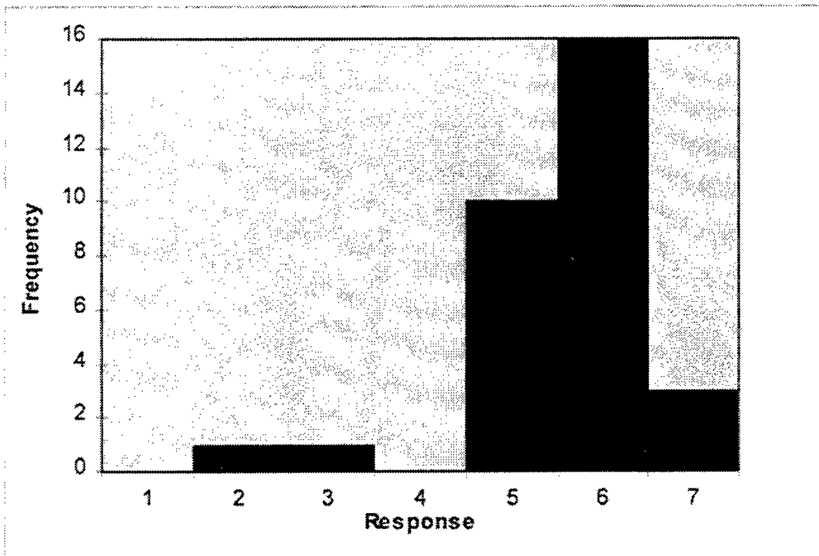
This appendix contains the results of the statistical analysis for hypothesis H_3, regarding decision confidence. Test question eight pertained to decision confidence (reference Chapter III or Appendix B). The response data used to conduct the analysis is shown in Appendix C, and the statistical analysis was accomplished using the software program Statistix®.

A table of Descriptive Statistics are provided, followed by histograms illustrating the distribution of responses for each treatment. Finally, the results of the Kruskal-Wallis test are shown. A discussion of these results is found in Chapter IV.

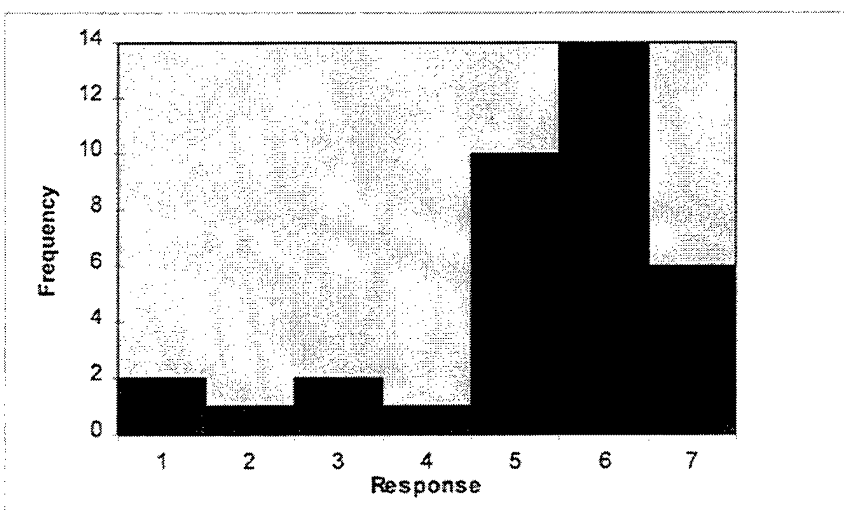
DESCRIPTIVE STATISTICS

	NONE_Q8	MED_Q8	HIGH_Q8
N	31	36	34
SUM	172	190	175
LO 95% CI	5.1715	4.7498	4.6029
MEAN	5.5484	5.2778	5.1471
UP 95% CI	5.9253	5.8057	5.6912
SD	1.0276	1.5604	1.5596
VARIANCE	1.0559	2.4349	2.4323
SE MEAN	0.1846	0.2601	0.2675
C.V.	18.520	29.566	30.300
MEDIAN	6.0000	6.0000	5.0000

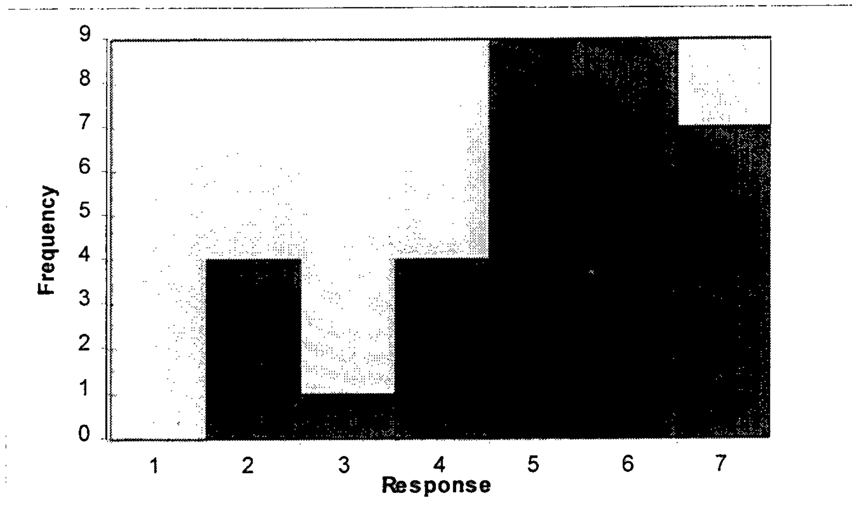
Distribution for Question 8 (No Extraneous Graphics)



Distribution for Question 8 (Medium Extraneous Graphics)



Distribution for Question 8 (High Extraneous Graphics)



KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR: NONE_Q7 MED_Q7 HIGH_Q7

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_Q7	49.1	31
MED_Q7	53.6	36
HIGH_Q7	50.0	34
TOTAL	51.0	101

KRUSKAL-WALLIS STATISTIC	0.4830
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.7855

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	386.540	193.270	0.24	0.7908
WITHIN	98	79645.0	812.704		
TOTAL	100	80031.5			

TOTAL NUMBER OF VALUES THAT WERE TIED 100
MAX. DIFF. ALLOWED BETWEEN TIES 0.00001

APPENDIX G. STATISTICAL ANALYSIS FOR H_4

This appendix contains the results of the statistical analysis for hypothesis H_4, regarding gender sensitivity analysis for hypotheses H_1, H_2, and H_3. The sensitivity analysis was accomplished by comparing the performance of male test subjects to that of female test subjects. The results of this analysis are summarized in Table 5 (information extraction accuracy) and Table 9 (decision accuracy and decision confidence), which are found in Chapter IV. The response data used to conduct the analysis is shown in Appendix C, and the statistical analysis was accomplished using the software program Statistix®.

Gender Sensitivity Analysis for H_1
(Test Questions 1 through 6)

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR NONE_M VS. NONE_F

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_M	92.1	126
NONE_F	96.4	60
TOTAL	93.5	186

KRUSKAL-WALLIS STATISTIC	1.0523
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.3050

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	719.421	719.421	1.05	0.3063
WITHIN	184	125761	683.481		
TOTAL	185	126480			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR MED_M VS. MED_F

VARIABLE	MEAN RANK	SAMPLE SIZE
MED_M	107.6	150
MED_F	110.6	66
TOTAL	108.5	216

KRUSKAL-WALLIS STATISTIC	0.3876
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.5336

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	415.505	415.505	0.39	0.5348
WITHIN	214	230056	1075.03		
TOTAL	215	230472			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR HIGH_M VS. HIGH_F

VARIABLE	MEAN RANK	SAMPLE SIZE
HIGH_M	95.7	114
HIGH_F	111.1	90
TOTAL	102.5	204

KRUSKAL-WALLIS STATISTIC	7.9536
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.0048

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	11819.3	11819.3	8.24	0.0045
WITHIN	202	289846	1434.88		
TOTAL	203	301665			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV BETWEEN TREATMENTS (MALE)

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_M	191.9	126
MED_M	196.2	150
HIGH_M	198.5	114
TOTAL	195.5	390

KRUSKAL-WALLIS STATISTIC	0.8749
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.6457

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	2724.72	1362.36	0.44	0.6526
WITHIN	387	1208713	3123.29		
TOTAL	389	1211438			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV BETWEEN TREATMENTS (FEMALE)

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_F	102.1	60
MED_F	102.6	66
HIGH_F	117.1	90
TOTAL	108.5	216

KRUSKAL-WALLIS STATISTIC	6.7213
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.0347

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	11418.5	5709.27	3.44	0.0332
WITHIN	213	353837	1661.21		
TOTAL	215	365256			

Gender Sensitivity Analysis for H_2
(Test Question 7)

DESCRIPTIVE STATISTICS

VARIABLE	N	MEAN	MEDIAN
HIGH7	34	5.5882	6.0000
HIGH7_F	15	5.2000	5.0000
HIGH7_M	19	5.8947	6.0000
MED7	36	5.6389	6.0000
MED7_F	11	5.7273	6.0000
MED7_M	25	5.6000	6.0000
NONE7	31	5.4516	6.0000
NONE7_F	10	5.2000	5.5000
NONE7_M	21	5.5714	6.0000

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR NONE7_M VS. NONE7_F

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE7_M	16.8	21
NONE7_F	14.3	10
TOTAL	16.0	31

KRUSKAL-WALLIS STATISTIC	0.5571
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.4554

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	42.6619	42.6619	0.55	0.4648
WITHIN	29	2254.84	77.7530		
TOTAL	30	2297.50			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV MED7_M VS. MED7_F

VARIABLE	MEAN RANK	SAMPLE SIZE
MED7_M	17.9	25
MED7_F	19.9	11
TOTAL	18.5	36

KRUSKAL-WALLIS STATISTIC	0.3068
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.5797

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	31.4509	31.4509	0.30	0.5870
WITHIN	34	3556.55	104.604		
TOTAL	35	3588.00			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR HIGH7_M VS. HIGH7_F

VARIABLE	MEAN RANK	SAMPLE SIZE
HIGH7_M	20.5	19
HIGH7_F	13.7	15
TOTAL	17.5	34

KRUSKAL-WALLIS STATISTIC 4.1862
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.0408

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	387.600	387.600	4.65	0.0387
WITHIN	32	2667.90	83.3719		
TOTAL	33	3055.50			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV BETWEEN TREATMENTS (MALE)

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE7_M	31.4	21
MED7_M	31.7	25
HIGH7_M	36.6	19
TOTAL	33.0	65

KRUSKAL-WALLIS STATISTIC 1.0607
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.5884

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	344.937	172.468	0.52	0.6012
WITHIN	62	20468.1	330.130		
TOTAL	64	20813.0			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR BETWEEN TREATMENTS (FEMALE)

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE7_F	17.8	10
MED7_F	21.8	11
HIGH7_F	16.6	15
TOTAL	18.5	36

KRUSKAL-WALLIS STATISTIC 1.6997
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.4275

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	177.593	88.7966	0.84	0.4398
WITHIN	33	3479.41	105.437		
TOTAL	35	3657.00			

Gender Sensitivity Analysis for H_3
(Test Question 8)

DESCRIPTIVE STATISTICS

VARIABLE	N	MEAN	MEDIAN
HIGH8	34	5.1471	5.0000
HIGH8_F	15	4.9333	5.0000
HIGH8_M	19	5.3158	6.0000
MED8	36	5.2778	6.0000
MED8_F	11	5.5455	6.0000
MED8_M	25	5.1600	6.0000
NONE8	31	5.5484	6.0000
NONE8_F	10	5.4000	6.0000
NONE8_M	21	5.6190	6.0000

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR NONE8_M VS. NONE8_F

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE8_M	16.2	21
NONE8_F	15.6	10
TOTAL	16.0	31

KRUSKAL-WALLIS STATISTIC 0.0345
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.8527

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	2.36190	2.36190	0.03	0.8563
WITHIN	29	2053.14	70.7979		
TOTAL	30	2055.50			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR MED8_M VS. MED8_F

VARIABLE	MEAN RANK	SAMPLE SIZE
MED8_M	17.2	25
MED8_F	21.4	11
TOTAL	18.5	36

KRUSKAL-WALLIS STATISTIC 1.2783
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.2582

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	129.895	129.895	1.29	0.2642
WITHIN	34	3426.61	100.783		
TOTAL	35	3556.50			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR HIGH8_M VS. HIGH8_F

VARIABLE	MEAN RANK	SAMPLE SIZE
HIGH8_M	18.8	19
HIGH8_F	15.9	15
TOTAL	17.5	34

KRUSKAL-WALLIS STATISTIC 0.7587
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.3837

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	71.6088	71.6088	0.75	0.3920
WITHIN	32	3042.89	95.0904		
TOTAL	33	3114.50			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV BETWEEN TREATMENTS (MALE)

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE8_M	35.2	21
MED8_M	30.7	25
HIGH8_M	33.6	19
TOTAL	33.0	65

KRUSKAL-WALLIS STATISTIC 0.7655
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.6820

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	238.412	119.206	0.38	0.6937
WITHIN	62	19694.1	317.647		
TOTAL	64	19932.5			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV BETWEEN TREATMENTS (FEMALE)

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE8_F	19.2	10
MED8_F	21.2	11
HIGH_F	16.1	15
TOTAL	18.5	36

KRUSKAL-WALLIS STATISTIC 1.6380
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.4409

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	172.830	86.4152	0.81	0.4535
WITHIN	33	3520.17	106.672		
TOTAL	35	3693.00			

APPENDIX H. STATISTICAL ANALYSIS FOR H_5

This appendix contains the results of the statistical analysis for hypothesis H_4, regarding training sensitivity analysis for hypotheses H_1, H_2, and H_3. The sensitivity analysis was accomplished by comparing the performance of test subjects with no prior training in graphical applications to those who have had some form of training (either formal or informal). The results of this analysis are summarized in Table 6 (information extraction accuracy) and Table 10 (decision accuracy and decision confidence), which are found in Chapter IV. The response data used to conduct the analysis is shown in Appendix C, and the statistical analysis was accomplished using the software program Statistix®.

Training Sensitivity Analysis for H_1
(Test Questions 1 through 6)

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR NONE_T VS. NONE_NT

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_T	91.1	132
NONE_NT	99.3	54
TOTAL	93.5	186

KRUSKAL-WALLIS STATISTIC	3.7154
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.0539

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	2540.12	2540.12	3.77	0.0537
WITHIN	184	123940	673.586		
TOTAL	185	126480			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR MED_T VS. MED_NT

VARIABLE	MEAN RANK	SAMPLE SIZE
MED_T	108.3	150
MED_NT	109.0	66
TOTAL	108.5	216

KRUSKAL-WALLIS STATISTIC	0.0183
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.8923

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	19.6364	19.6364	0.02	0.8927
WITHIN	214	230452	1076.88		
TOTAL	215	230472			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR HIGH_T VS. HIGH_NT

VARIABLE	MEAN RANK	SAMPLE SIZE
HIGH_T	95.7	114
HIGH_NT	111.1	90
TOTAL	102.5	204

KRUSKAL-WALLIS STATISTIC	7.9536
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.0048

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	11819.3	11819.3	8.24	0.0045
WITHIN	202	289846	1434.88		
TOTAL	203	301665			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV BETWEEN TREATMENTS (TRAINING)

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_T	193.0	132
MED_T	200.8	150
HIGH_T	201.8	114
TOTAL	198.5	396

KRUSKAL-WALLIS STATISTIC	1.9136
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.3841

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	6059.84	3029.92	0.96	0.3871
WITHIN	393	1244805	3167.44		
TOTAL	395	1250865			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV BETWEEN TREATMENTS (NO TRAINING)

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE_NT	102.1	54
MED_NT	97.6	66
HIGH_NT	113.3	90
TOTAL	105.5	210

KRUSKAL-WALLIS STATISTIC	6.2397
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.0442

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	10244.4	5122.20	3.19	0.0423
WITHIN	207	332896	1608.19		
TOTAL	209	343140			

Training Sensitivity Analysis for H_2 (Test Question 7)

DESCRIPTIVE STATISTICS

VARIABLE	N	MEAN	MEDIAN
HIGH7	34	5.5882	6.0000
HIGH_NT	15	5.2667	5.0000
HIGH7_T	19	5.8421	6.0000
MED7	36	5.6389	6.0000
MED_NT	11	5.9091	6.0000
MED7_T	25	5.5200	6.0000
NONE7	31	5.4516	6.0000
NONE7_NT	9	4.8889	5.0000
NONE_T	22	5.6818	6.0000

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR NONE7_T VS. NONE7_NT

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE7_T	17.0	22
NONE7_NT	13.4	9
TOTAL	16.0	31

KRUSKAL-WALLIS STATISTIC	1.0815
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.2984

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	82.8232	82.8232	1.08	0.3063
WITHIN	29	2214.68	76.3682		
TOTAL	30	2297.50			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR MED7_T VS. MED_7_NT

VARIABLE	MEAN RANK	SAMPLE SIZE
MED7_T	18.3	25
MED7_NT	18.9	11
TOTAL	18.5	36

KRUSKAL-WALLIS STATISTIC	0.0204
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.8863

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	2.09455	2.09455	0.02	0.8888
WITHIN	34	3585.91	105.468		
TOTAL	35	3588.00			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV HIGH_T VS. HIGH_NT

VARIABLE	MEAN RANK	SAMPLE SIZE
HIGH_T	20.0	19
HIGH_NT	14.3	15
TOTAL	17.5	34

KRUSKAL-WALLIS STATISTIC 2.9686
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.0849

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	274.863	274.863	3.16	0.0848
WITHIN	32	2780.64	86.8949		
TOTAL	33	3055.50			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV BETWEEN TREATMENTS (TRAINING)

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE7_T	32.0	22
MED7_T	33.0	25
HIGH7_T	35.9	19
TOTAL	33.5	66

KRUSKAL-WALLIS STATISTIC 0.4857
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.7844

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	164.711	82.3553	0.24	0.7914
WITHIN	63	21878.8	347.282		
TOTAL	65	22043.5			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV BETWEEN TREATMENTS (NO TRAINING)

VARIABLE	MEAN RANK	SAMPLE SIZE
LQ7_NT	15.9	9
MQ7_NT	22.2	11
HQ7_NT	16.2	15
TOTAL	18.0	35

KRUSKAL-WALLIS STATISTIC 2.9291
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.2312

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	287.096	143.548	1.51	0.2366
WITHIN	32	3045.40	95.1689		
TOTAL	34	3332.50			

Training Sensitivity Analysis for H_3
(Test Question 8)

DESCRIPTIVE STATISTICS

VARIABLE	N	MEAN	MEDIAN
HIGH8	34	5.1471	5.0000
HIGH8_NT	15	5.0000	5.0000
HIGH8_T	19	5.2632	6.0000
MED8	36	5.2778	6.0000
MED8_NT	11	5.8182	6.0000
MED8_T	25	5.0400	5.0000
NONE8	31	5.5484	6.0000
NONE8_NT	9	5.4444	6.0000
NONE8_T	22	5.5909	6.0000

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR NONE8_T VS. NONE8_NT

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE8_T	15.6	22
NONE8_NT	16.9	9
TOTAL	16.0	31

KRUSKAL-WALLIS STATISTIC	0.1462
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.7022

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	10.0202	10.0202	0.14	0.7090
WITHIN	29	2045.48	70.5338		
TOTAL	30	2055.50			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV FOR MED8_T VS. MED8_NT

VARIABLE	MEAN RANK	SAMPLE SIZE
MED8_T	17.3	25
MED8_NT	21.1	11
TOTAL	18.5	36

KRUSKAL-WALLIS STATISTIC	1.0835
P-VALUE, USING CHI-SQUARED APPROXIMATION	0.2979

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	110.095	110.095	1.09	0.3047
WITHIN	34	3446.41	101.365		
TOTAL	35	3556.50			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV HIGH8_T VS. HIGH8_NT

VARIABLE	MEAN RANK	SAMPLE SIZE
HIGH8_T	19.1	19
HIGH8_NT	15.5	15
TOTAL	17.5	34

KRUSKAL-WALLIS STATISTIC 1.1759
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.2782

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	1	110.977	110.977	1.18	0.2850
WITHIN	32	3003.52	93.8601		
TOTAL	33	3114.50			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV BETWEEN TREATMENTS (TRAINING)

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE8_T	34.7	22
MED8_T	31.3	25
HIGH8_T	35.1	19
TOTAL	33.5	66

KRUSKAL-WALLIS STATISTIC 0.5913
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.7440

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	203.536	101.768	0.29	0.7533
WITHIN	63	22170.5	351.912		
TOTAL	65	22374.0			

KRUSKAL-WALLIS ONE-WAY NONPARAMETRIC AOV BETWEEN TREATMENTS (NO TRAINING)

VARIABLE	MEAN RANK	SAMPLE SIZE
NONE8_NT	20.4	9
MED8_NT	21.6	11
HIGH8_NT	13.9	15
TOTAL	18.0	35

KRUSKAL-WALLIS STATISTIC 5.0010
P-VALUE, USING CHI-SQUARED APPROXIMATION 0.0820

PARAMETRIC AOV APPLIED TO RANKS

SOURCE	DF	SS	MS	F	P
BETWEEN	2	441.269	220.634	2.76	0.0784
WITHIN	32	2558.73	79.9604		
TOTAL	34	3000.00			

BIBLIOGRAPHY

Abdolmohammadi, M. and A. Wright. "An Examination of the Effects of Experience and Task Complexity on Audit Judgements," *The Accounting Review*, January 1987: 1-13.

Apple Publishing. The Basic Elements of Design: *Guidelines for Creating Professional-Looking Documents With Your Laser Writer*. Cupertino CA: Apple Computer, Inc., 1991.

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, Maxine D. "Mainframe Business Graphics," *Datamation,* May 1, 1984: 89-95.

Butler, Janet B. and R. David Mautz, Jr. "Multimedia presentations and learning: A laboratory experiment," *Issues in Accounting Education*, *11* Num. 2, 1996: 261-280.

Campbell, Donald T. and Julian C. Stanley. *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand McNally College Publishing Company, 1963.

Christensen, David S. and Albert Larkin. "Criteria for High Integrity Graphics," *Journal of Managerial Issues*, *IV*, Num. 1, Spring 1992: 130-153.

Claris Publishing. *ClarisImpact User's Guide*. Santa Clara CA: Claris Corporation, 1993.

Cleveland, William S. *The Elements of Graphing Data*. Monterey: Wadsworth Advanced Books and Software, 1985.

Conover, W. J. *Practical NonParametric Statistics* (2^{nd} Edition). John Wiley & Sons: New York, 1971.

Cooper, Donald R. and C. William Emory. *Business Research Methods* (5^{th} Edition). Chicago: Irwin, 1995.

Corbone, Robert and Wilpen L. Gorr. "Accuracy of Judgmental Forecasting of Time Series," *Decision Sciences, 16*, Spring 1985: 153-160.

Davis, Larry R. "Report Format and the Decision Maker's Task: An Experimental Investigation," *Accounting, Organizations and Society, 14*, Num. 5/6, 1989: 495-508.

Decisioneering. *Crystal Ball: Forecasting and Risk Analysis for Spreadsheet Users.* Denver: Decisioneering, Inc., 1993.

DeSanctis, Geradine. "Computer Graphics as Decision Aids: Direction for Research," *Decision Sciences, 15*, Fall 1984: 463-385.

DeSanctis, Geradine and Sirkka L. Jarvenpaa. 1989. Graphical presentation of accounting data for financial forecasting: An experimental investigation. *Accounting Organizations and Society, 14* (5/6): 509-525.

Devore, Jay L. *Probability and Statistics for Engineering and the Sciences* (4th Edition). Albany NY: Duxbury Press, 1995.

Dickson, Gary W., Geradine DeSanctis, and D. J. McBride. "Understanding the Effectiveness of Computer Graphics for Decision Support: A Cumulative Experimental Approach," *Communications of the ACM, 29*, Num. 21, January 1986: 40-47.

Horton, William. *Illustrating Computer Documentation: The Art of Presenting Information Graphically on Paper and Online*. New York: John Wiley & Sons, Inc., 1991.

Hwang, Mark I.H. and Bruce J. P. Wu. "The Effectiveness of Computer Graphics for Decision Support: A Meta-Analytical Integration of Research Finding," *Database,* Fall, 1990: 11-20.

Jarvenpaa, Sirkka L. and Gary W. Dickson. "Myths vs. Facts About Graphics in Decision Making," *Spectrum, 3*, Num. 1, February, 1986: 1-4.

Larkin, Albert A. "Misleading Graphics: Can Decision Makers be Affected by Their Use?" MS Thesis, AFIT/GCA/LSY/90S-18. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1990 (AD-A229621).

Latin, Anita E. and Anthony Villanueva. "The Effect of Three-Dimensional Graphs on Decision Making," MS Thesis, AFIT/GSM/LAS/94S-7. School of Logistics and Acquisition Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1994 (AD-A285034).

MacKay, David B. and Angelina Villareal. "Performance Differences in the Use of Graphic and Tabular Displays of Multivariate Data," *Decision Sciences, 18*, 1987: 535-547.

Microsoft. *Getting Results With Microsoft Office for Windows 95*. Microsoft Corporation, 1995.

Montazemi, Ali R. "Comments on 'A Meta-Analysis of the Relative Effects of Tabular and Graphic Display Formats on Decision-Making Performance'," *Human Performance, 4*, Num. 2, 1991: 127-146.

Needleman, Ted. Great Presentation: "Taking Center Stage; Use Your PC to Sell Your Point of View," *Accounting Technology*, *4*, June 1993: 14-25.

Neter, John and William Wasserman. *Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs*. Homewood IL: Irwin, 1974.

Peeler, David L., Jr. "Background graphics: A decision-making hindrance or enhancement?" MS Thesis, AFIT/GCA/LAS/96S-12. School of Logistics and Acquisition Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, 1996 (AD-A319517).

Schaubroeck, John and Krishnamurty Muralidhar. "A Meta-Analysis of the Relative Effects of Tabular and Graphic Display Formats on Decision-Making Performance," *Human Performance, 4*, Num. 2, 1991: 147-154.

Tufte, Edward R. *The Visual Display of Quantitative Information*. Cheshire CT: Graphics Press, 1983.

Tullis, T. S. "An Evaluation of Alphanumeric, Graphic, and Color Display," *Human Factors*, *25*, Num. 5, 1981: 541-550.

Ward, Douglas B. "The Effectiveness of Sidebar Graphics," *Journalism Quarterly, 69*, Issue 2, Summer 1992: 318-328.

Zmud, R.W., E. Blocher, and R. P. Moffie. "The Impact of Color Graphic Report Formats on Decision Performance and Learning," *Proceedings of the Fourth International Conference on Information Systems*. 1983: 179-193.

VITA

Captain Thomas S. Tingley earned a Bachelor of Science degree in Electrical Engineering from the United States Air Force Academy in 1990. Following graduation, he was assigned to the National Aerospace Intelligence Center (NAIC) at Wright-Patterson AFB, Ohio. At NAIC, he served as a foreign systems engineering analyst in the Command, Control, and Communication (C^3) analysis division. During his assignment at Wright-Patterson, Captain Tingley also participated in the All-Air Force wrestling team. Representing the Air Force, he placed in the top eight at the United States National Greco-Roman Wrestling Championships, every year between 1992 and 1994. He also finished sixth at the 1994 United States World Team Trials.

In 1994, Captain Tingley was jointly assigned to Headquarters, North American Aerospace Defense Command (NORAD) and Headquarters, United States Space Command (USSPACECOM) at Peterson AFB, Colorado. While at NORAD/USSPACECOM, he served in the J6 Directorate as Network Manager for the Cheyenne Mountain Operations Center - Command Automation System. In 1996, he was assigned to the Air Force Institute of Technology (AFIT) at Wright-Patterson AFB, Ohio, to pursue a Masters degree in Systems Management. Upon graduation from AFIT in September 1997, he will be assigned to Headquarters, Air Force Operational Test and Evaluation Center (AFOTEC), at Kirtland AFB, New Mexico.

Permanent Address: 2938 Sandy Lake/Grove City Road
Stoneboro, PA 16153

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1997	Master's Thesis

4. TITLE AND SUBTITLE

THE EFFECTS OF EXTRANEOUS PRESENTATION GRAPHICS ON DECISION MAKING

5. FUNDING NUMBERS

6. AUTHOR(S)

Thomas S. Tingley, Captain, USAF

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)

Air Force Institute of Technology
2950 P Street
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GSM/LAS/97S-3

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

HQ AFOTEC/XRE
Lt Col Jancik
8500 Gibson Blvd SE
Kirtland AFB, NM 87117-5558

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)*

Almost every professional presentation includes some form of graphics. Many previous researchers have analyzed the relative effectiveness of graphical decision-aids, such as pie charts or bar graphs. Very little research, however, has been conducted to determine the effects of extraneous graphics, such as pictures and logos, on decision-making. This study employed an experiment to determine whether extraneous graphics hinder or enhance a decision-maker's ability to obtain the pertinent information from a professional presentation. Other factors such as gender and graphics training were also analyzed for any influence on the results obtained. The results of the study indicated a statistically significant effect on information extraction accuracy due to the use of extraneous graphics, but an effect was not discovered regarding decision accuracy, or decision confidence. Information extraction accuracy was found to be inversely related to the level of extraneous graphics employed in a presentation. Additionally, the effect on extraction accuracy was found to be more predominant for females and for people without prior experience or training (either formal or informal) in graphics applications. The tests also showed that most people prefer a "medium" amount of extraneous graphics rather than a large amount or no extraneous graphics.

14. Subject Terms

Graphics, Computer Graphics, Decision Making, Decision Aids

15. NUMBER OF PAGES 150

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

AFIT Control Number AFIT/GSM/LAS/97S-3

AFIT RESEARCH ASSESSMENT

The purpose of this questionnaire is to determine the potential for current and future applications of AFIT thesis research. **Please return completed questionnaire** to: AIR FORCE INSTITUTE OF TECHNOLOGY/LAC, 2950 P STREET, WRIGHT-PATTERSON AFB OH 45433-7765. Your response is **important.** Thank you.

1. Did this research contribute to a current research project? a. Yes b. No

2. Do you believe this research topic is significant enough that it would have been researched (or contracted) by your organization or another agency if AFIT had not researched it?
a. Yes b. No

3. **Please estimate** what this research would have cost in terms of manpower and dollars if it had been accomplished under contract or if it had been done in-house.

Man Years__________ $__________

4. Whether or not you were able to establish an equivalent value for this research (in Question 3), what is your estimate of its significance?

a. Highly Significant b. Significant c. Slightly Significant d. Of No Significance

5. Comments (Please feel free to use a separate sheet for more detailed answers and include it with this form):

____
Name and Grade

____
Organization

____
Position or Title

____
Address