An examination of factors affecting audit planning across accounts

Usoff, Catherine Anne, Ph.D.

The Ohio State University, 1994
AN EXAMINATION OF FACTORS AFFECTING AUDIT PLANNING ACROSS ACCOUNTS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By

Catherine Anne Usoff, B.S., M.B.A.

* * * * *

The Ohio State University
1994

Dissertation Committee:

E. E. Spires
W. I. Notz
T. E. Nygren
D. E. Wallin

Approved by:

E. E. Spires
Advisor
Department of Accounting and Management of Information Systems
To my husband, Joseph and our daughter, Rebecca
ACKNOWLEDGMENTS

I offer sincere appreciation to my advisor, Dr. Eric E. Spires for his guidance and for providing me with an excellent role model. I appreciate the time and effort of the other members of my dissertation committee, Drs. David E. Wallin, William I. Notz and Thomas E. Nygren. Much gratitude is expressed to the auditors who participated in my research and to my fellow Ph.D. students who assisted in the design of the materials. I am grateful for comments received at workshops at The Ohio State University, Bentley College, Boston College, the University of Massachusetts at Boston, and the Northeast Regional AAA Doctoral Consortium. Special thanks to Dr. Arnold Wright for his many insightful comments and suggestions.

To my friend, Chris Hogan, thanks for keeping me sane (well, you tried anyway) and for making a Ph.D. program even fun at times.

To my husband, Joe, I am truly thankful for your love and your faith in me. To our daughter, Rebecca, thank you for being a great kid.
VITA

December 4, 1960 ............................................. Born - Quincy, Massachusetts

1982 .......................................................... B.S., Boston College, Chestnut Hill, Massachusetts


1985-1986 .......................................................... Graduate Teaching Associate, The Ohio State University, Columbus, Ohio

1986 .......................................................... M.B.A., The Ohio State University, Columbus, Ohio

1987-1989 .......................................................... Instructor, New Mexico State University, Alamogordo, New Mexico

1989-1993 .......................................................... Graduate Teaching Associate, The Ohio State University, Columbus, Ohio

1993-present ..................................................... Assistant Professor, Bentley College, Waltham, Massachusetts

FIELDS OF STUDY

Major Field: Accounting and Management of Information Systems

Minor Fields: Applied Statistics

Quantitative Psychology
TABLE OF CONTENTS

DEDICATION ................................................................. ii
ACKNOWLEDGMENTS .................................................. iii
VITA ........................................................................ iv
LIST OF TABLES .......................................................... x
LIST OF FIGURES ......................................................... xi

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Description of the Research</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Expected Contributions of this Research</td>
<td>9</td>
</tr>
<tr>
<td>1.3.1 Provides Link between Theoretical and Experimental Audit Planning Research</td>
<td>10</td>
</tr>
<tr>
<td>1.3.2 Provides Experimental Evidence Regarding Substantive Testing Across Accounts</td>
<td>10</td>
</tr>
<tr>
<td>1.3.3 Provides Evidence about Auditors' Priorities in Planning an Audit</td>
<td>11</td>
</tr>
<tr>
<td>1.4 Organization of the Dissertation</td>
<td>11</td>
</tr>
<tr>
<td>II. LITERATURE REVIEW</td>
<td>12</td>
</tr>
<tr>
<td>2.1 Audit Planning Research</td>
<td>13</td>
</tr>
<tr>
<td>2.1.1 Experimental Audit Planning Research</td>
<td>13</td>
</tr>
<tr>
<td>2.1.2 Archival Audit Planning Research</td>
<td>20</td>
</tr>
</tbody>
</table>
2.2 Theoretical Research

2.2.1 Decision-Theoretic Research

2.2.2 Game Theory Research

2.3 Archival Research on the Incidence and Consequence of Errors and Irregularities

2.4 Auditors' Assessments of Consequences

2.5 Summary of Relevant Literature

III. DEVELOPMENT OF RESEARCH HYPOTHESES

3.1 Effect of Probability and Consequence Levels on Planning Judgments

3.1.1 Decision-Theoretic Model

3.1.1.1 Assumptions

3.1.1.2 The Auditor's Choice as a Minimization of Expected Cost

3.1.1.3 Numerical Example

3.1.2 Statement of Research Hypotheses

3.2 Effect of the Auditor's Type on Substantive Testing Levels

3.2.1 CJM Scaling and the Effect of Type

3.3 Alternative Dependent Variables

IV. RESEARCH APPROACH

4.1 Suitability of the Experimental Approach to Audit Planning Research

4.2 Suitability of Conjoint Measurement Approach

4.3 Suitability of Multiple Regression Analysis of the SAC Task

V. DEVELOPMENT AND DESCRIPTION OF RESEARCH

5.1 Overview of the Study

5.1.1 Consulting Auditors
5.1.2 Administration of the Study and Research Design

5.2 CJM Task
5.2.1 Development of the CJM Task
5.2.2 Pilot Testing of the CJM Task
5.2.3 Administration of the CJM Task
5.2.4 CJM Analysis
5.2.4.1 Previous Applications of CJM in Accounting Research
5.2.4.2 CMSCAL Applied to Audit Planning Judgments
5.2.5 Examples of Trade-off Violations of Joint Independence Axioms
5.2.5.1 Probability and Effort
5.2.5.2 Consequence and Effort
5.2.5.3 Probability and Consequence

5.3 The SAC Task and Questionnaire
5.3.1 Development of the SAC Task
5.3.1.1 Background Information
5.3.1.2 Designing Levels of the Probability Factor
5.3.1.3 Designing Levels of the Consequence Factor
5.3.2 Dependent Variables
5.3.3 Post-Experimental Questionnaire
5.3.4 Pilot Testing of the SAC Task

VI RESULTS
6.1 Descriptive Data
6.2 Results of Manipulation Checks ........................................ 72
  6.2.1 Manipulation Checks Related to Consequence Levels 72
  6.2.2 Manipulation Checks Related to Probability Levels . 75
  6.2.3 Summary of Results Related to Manipulation Checks 77
6.3 Regression Model for Planning Judgments ................. 78
6.4 Results Related to Tests of Hypothesis One .......... 83
  6.4.1 Reduced Probability Models for AR ........... 84
  6.4.2 Reduced Probability Models for PPE ......... 85
  6.4.3 Summary of the Effect of Probability ............ 87
6.5 Results of Tests of Hypothesis Two ..................... 87
  6.5.1 Reduced Consequence Models for AR ................. 88
  6.5.2 Reduced Consequence Models for PPE .............. 92
  6.5.3 Summary of the Effect of Consequence ............... 95
6.6 Results of Tests of Hypothesis Three .................... 97
  6.6.1 Reduced Models for the Probability by Consequence Interaction ............................................................ 97
6.7 Results of Tests of Hypotheses Four and Five .......... 99
  6.7.1 Reduced Models for the TypeP Variable ........... 100
  6.7.2 Reduced Models for the TypeC Variable ............ 102
6.8 Results of Tests of Hypothesis Six ...................... 105
6.9 Importance of Factors and Trade-offs Made between Factors . 108
7.1 Results Associated with the Major Objective of this Study .. 110
  7.1.1 Implications of Results Associated with the Major Objective ......................................................... 111
7.2 Link between Theoretical and Experimental Audit Planning Research ........................................ 112
7.3 Experimental Evidence on More than One Account within an Audit ........................................... 113
7.4 Evidence about Auditors' Priorities ......................................................... 115
7.5 Conclusion ......................................................................................... 116
    7.5.1 Implications for Audit Planning Research ............ 116
    7.5.2 Implications for Audit Practice .......................... 117

APPENDICES

A. CJM Task ................................................................. 118
B. CJM Instructions .......................................................... 121
C. Version A of SAC Task .................................................. 124
D. Introductory Sheet Given to All Subjects ................. 145
E. Description of the Axioms of CJM ............................... 147
F. Simple Polynomial Models Using Three Proposed Factors . 152

REFERENCES ....................................................................................... 156
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Maximum Acceptable Cost of Decreasing Detection Risk by One Percent</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Prototype Rankings</td>
<td>39</td>
</tr>
<tr>
<td>2.</td>
<td>Score Values for Combinations of Factors</td>
<td>43</td>
</tr>
<tr>
<td>3.</td>
<td>Participating Auditors, Descriptive Data</td>
<td>44</td>
</tr>
<tr>
<td>5.</td>
<td>Results of Manipulation Checks for Consequence</td>
<td>71</td>
</tr>
<tr>
<td>6.</td>
<td>Results of Manipulation Checks for Probability</td>
<td>74</td>
</tr>
<tr>
<td>7.</td>
<td>Full Regression Models - Results</td>
<td>76</td>
</tr>
<tr>
<td>8.</td>
<td>Reduced Probability Models for AR - Results</td>
<td>81</td>
</tr>
<tr>
<td>9.</td>
<td>Reduced Probability Models for PPE - Results</td>
<td>84</td>
</tr>
<tr>
<td>10.</td>
<td>Reduced Consequence Models for AR - Results</td>
<td>86</td>
</tr>
<tr>
<td>11.</td>
<td>Reduced Consequence Models for PPE - Results</td>
<td>90</td>
</tr>
<tr>
<td>12.</td>
<td>Pearson Correlations between Probability and Consequence</td>
<td>93</td>
</tr>
<tr>
<td>13.</td>
<td>Reduced Regression Models for the Probability by Average Consequence Interaction - Results</td>
<td>96</td>
</tr>
<tr>
<td>14.</td>
<td>Reduced Regression Models for TypeP - Results</td>
<td>101</td>
</tr>
<tr>
<td>15.</td>
<td>Reduced Regression Models for TypeC - Results</td>
<td>103</td>
</tr>
<tr>
<td>16.</td>
<td>Pearson Correlations between Dependent Variables</td>
<td>107</td>
</tr>
<tr>
<td>17.</td>
<td>Most Important Factor</td>
<td>108</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>DESCRIPTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Overview of this Research and how it Relates to Previous Research</td>
<td>5</td>
</tr>
<tr>
<td>2.</td>
<td>Cost Matrix for Level of Substantive Testing Choice</td>
<td>37</td>
</tr>
<tr>
<td>3.</td>
<td>Overall Experimental Design</td>
<td>52</td>
</tr>
<tr>
<td>4.</td>
<td>Design of SAC Task</td>
<td>72</td>
</tr>
<tr>
<td>5.</td>
<td>Full Regression Model and Definitions of Variables</td>
<td>79</td>
</tr>
<tr>
<td>6.</td>
<td>Description of Factor Levels</td>
<td>118</td>
</tr>
<tr>
<td>7.</td>
<td>Sample of Card Used in CJM Task</td>
<td>120</td>
</tr>
<tr>
<td>8.</td>
<td>Flowchart for the Diagnosis of Simple Polynomials</td>
<td>151</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

1.1 Background

When performing an audit, certified public accountants (CPAs) are required to follow the Statements on Auditing Standards (SASs) issued by the American Institute of Certified Public Accountants (AICPA). The first standard of field work prescribed in the SASs is that the audit should be adequately planned and properly supervised. This standard provides general guidance by indicating some of what the auditor should consider when planning the audit. One of the items specified is "conditions that may require extension or modification of audit tests, such as the risk of material errors or irregularities or the existence of related party transactions" (AICPA 1989, Section 311.03). This implies that more testing should be planned if there is a higher risk of material errors or irregularities. Other planning guidelines appear throughout the statements. For example, SAS No. 47, "Audit Risk and Materiality in Conducting an Audit" states that in planning the audit, the auditor should consider both audit risk and materiality.

Audit risk is defined as "the risk that the auditor may unknowingly fail to appropriately modify his opinion on financial statements that are materially misstated" (AICPA 1989, Section 312.02). This risk consists of three components; inherent risk (IR), control risk (CR) and detection risk (DR). IR refers to the probability that there is a material misstatement. A misstatement may be unintentional (an error) or intentional (an
irregularity)\(^1\). CR refers to the probability that the system of internal control will not prevent or detect a material misstatement. DR refers to the probability that audit tests will not result in the detection of a material misstatement. IR and CR are the result of characteristics of a client's accounts and systems. DR is a result of the audit procedures planned and performed by the auditor. A literal application of the audit risk model would consist of the auditor first estimating the level of IR and CR. He or she would then plan an appropriate level of audit procedures to result in a DR level such that when combined with the existing IR and CR, the overall audit risk will be acceptable.

Materiality is defined as "the magnitude of an omission or misstatement of accounting information that, in light of surrounding circumstances, makes it probable that the judgment of a reasonable person relying on the information would have been changed or influenced by the omission or misstatement" (AICPA 1989, Section 312.06).\(^2\)

Whether the auditor applies the audit risk model literally or not, he or she is expected to plan the audit so that enough evidence is gathered to provide an acceptably low level of risk that the financial statements are materially misstated. The auditor must estimate in the planning stage what the risk of misstatement is in order to plan sufficient testing. SAS 53, "Auditor's Responsibility to Detect and Report Errors and Irregularities" provides some guidance by listing factors that, if present, increase the risk of misstatement. The AICPA recognizes that it is not possible or desirable to ensure that there is not material misstatement. In that vein, SAS 53 states that "the subsequent discovery that a material misstatement exists in the financial statements does not, in and

\(^1\) Irregularities consist of fraud and defalcations. Fraud involves misstating an amount or item to make the financial statements misleading. Defalcation refers to stealing an asset.

of itself, evidence inadequate planning, performance or judgment on the part of the auditor" (AICPA 1989, Section 316.08).

The AICPA's position does not appear to protect the welfare of an audit firm when an error or irregularity is discovered after the audit report is issued. Even if an audit firm performs a high quality audit; i.e. maintains low audit risk, the firm is subject to business risk due to its role in the capital market. When a client company fails or is found to have management fraud, the injured parties may sue the company's auditor. Despite what the auditing standards recognize as a limitation of an audit, the court system appears to hold the auditor responsible for detecting fraud (Palmrose 1987). In light of the gap between what the public expects from auditors and what the profession maintains as the responsibility of the auditor, it may be that business risk is a major consideration when planning an audit; not because it is required by the auditing standards but because it is a necessity of the environment in which auditors operate. This research is intended as an exploration of what auditors consider when planning an audit and how important various factors are to individual auditors.

1.2 Description of the Research

This research investigates how three factors - probability, effort and consequence - affect the auditor's preference for relative levels of substantive testing across accounts. "Probability" refers to the likelihood that there is a material error or irregularity in an

---

3 Business risk refers to the risk that the firm will suffer a financial loss.

4 Substantive testing consists of audit procedures directed at ascertaining whether the balances reported in the accounts are fairly stated. Usually the level of substantive testing depends, in part, on how much reliance is placed on the client's internal control system. Since differences in auditors' assessments of the internal control system would unnecessarily add to the complexity of this study, this research assumes controls will not be relied on and focuses on substantive tests only. In support of this assumption, Mock and Wright (1993) found that controls were not relied on in 80.5% of the audits they studied.

5 The term "error" will be used throughout the remainder of the dissertation to represent an error or irregularity.
account. "Effort" refers to the amount of audit effort that would be necessary to increase the level of substantive testing to a high level. "Consequence" refers to the expected costs to the auditor and the audit firm if there is a material error in an account and it is not detected in the audit.

Based on the auditing standards, previous archival and experimental audit planning research has focused on only one of these factors - probability of material error in an account. This factor is equivalent to a combined inherent and control risk assessment (as defined in section 1.1) and therefore it is expected that the level of planned audit testing would be associated with this factor. This stream of research, however, has not been able to establish a clear link between the level of audit testing (measured by the number of audit hours planned for an account) and the assessed probability of error in the account (Bedard and Wright 1992, Kaplan and Reckers 1989, and Mock and Wright 1993).

The major objective of this research is to explore potential explanations for the lack of significant association between the probability of error and planned audit hours in previous audit planning research. This dissertation explores three potential explanations for the lack of a significant association found in the previous research. One potential explanation is that another factor, namely consequence, is important to auditors' planning judgments. A second possible explanation relates to individual differences between auditors' consideration of two factors, probability and consequence. The third potential explanation is that the traditional dependent variable used in this line of research, planned audit hours for an account, is not a reasonable measure to use to compare the auditors' perceptions of the audit intensity required for the account. Figure 1 illustrates how the potential explanations relate to the audit planning process.
Representation of Relationship Tested in Previous Research

Given a particular audit situation, the auditors were expected to:

- Assess probability of error in an account
- Indicate number of audit hours for the account

Representation of Relationships Tested in this Research

Given a particular audit situation, the auditors are expected to:

- Assess probability of error in an account
- Assess consequence of not finding an error in an account
- Indicate relative level of substantive testing for account (Based on acceptable detection risk)
- Indicate number of audit hours for the account
- Variability introduced by types of tests, extent of tests and levels of audit personnel

Weights (depend on type?)

Figure 1. Overview of this Research and how it Relates to Previous Research
Support for considering the first explanation comes from the theoretical auditing research. In contrast to experimental audit planning research, the theoretical research related to audit planning typically includes, in some form, all three of the factors proposed here - probability, effort and consequence. The results of the decision-theoretic (Kinney 1975 and Menzefricke 1984) and game theory approaches (Matsumura and Tucker 1992 and Shibano 1990) suggest that consequences associated with not finding an error would have a significant impact on the auditor's planning decisions and the resulting level of audit risk. If consequence is as important to audit planning judgments as suggested by the theoretical research, the omission of this variable in previous empirical and experimental audit planning research could help to explain a lack of significant association between probability and audit hours. Each previous experiment used only one account with a between subjects manipulation of the probability of error. If the participating auditors considered the consequence associated with not finding an error in that account, the probability manipulation may not have had the intended effect on audit hours. For example, if accounts receivable was the account presented and an auditor considered there to be a high consequence associated with not finding an existing error in accounts receivable, he or she may have planned a substantial number of audit hours regardless of the perceived probability of error. The expected differences in audit hours due to different levels of probability of error may not have materialized because the auditors focused more on the consequence level.

The second explanation is suggested by the great amount of variability found by other researchers (Bedard and Wright 1992, Joyce 1976, and Kaplan and Reckers 1984) between auditors making planning judgments. It appears that individual differences have a strong effect on the results of this research. In an attempt to explain some of the
variability, an element of individual difference is incorporated into the regression models used in this study. A "type" variable is included which indicates the factor that the auditor considers to be the most important.

Support for suspecting a dependent variable measurement problem (the third potential explanation) comes from Joyce (1976) who used a within subjects design to study auditors' planning decisions. His subjects displayed very different ranges of the number of audit hours they would plan over the 20 or 36 cases that each was given. Typically, experimental audit planning research uses a between subjects design with each subject receiving one case. The number of audit hours planned by a subject for one case is difficult to interpret because we do not know how the subject's reaction to this particular situation compares to how he or she would react to other situations. It is possible that two subjects could react in the same way, both wanting to plan an average level of hours for the client but have different approaches to testing so the number of audit hours would not reflect the similar intentions. The present research adds two additional dependent variables to the traditional planned audit hours to attempt to identify intentions and reduce the effect of individual variability inherent in the number of audit hours planned. Subjects are asked to indicate a "level of concern," as well as a "relative level of substantive testing" in addition to the number of audit hours. The three dependent variables are expected to be positively associated with each other. The level of concern and relative level of substantive testing, however, are expected to have more shared variability with probability of error than does the number of audit hours. The number of audit hours is expected to have more unexplained variability due to individual test planning differences between auditors.
The research approach that was used to explore these explanations was experimental, consisting of two tasks and a questionnaire. One task that was given to auditor subjects, the Conjoint Measurement Task, required that they rank 27 combinations. Each combination consisted of a level of probability (high, medium, or low), a level of effort\(^6\) (low, medium, or high) and a level of consequence (high, medium, or low). The rankings represented the auditor's relative preference for increasing substantive testing on an account having the characteristics described by each combination. Subjects' responses (rankings) were analyzed using a conjoint measurement (CJM) algorithm. CJM analysis indicated the relative importance placed on the factors, any trade-offs auditors made between the factors and whether a simple polynomial model is an appropriate representation of how factors were combined. This task addressed the first potential explanation by examining each auditor's relative weights on probability and consequence.

The factor that each subject considered to be the most important was used to classify him or her as a probability or consequence "type." The analysis of subjects' rankings indicate that forty percent of the 65 subjects thought probability was the most important factor and fifty-two percent thought consequence was the most important factor.\(^7\) Approximately one third of the subjects (21 subjects, 32 percent) ranked the 27 combinations in a way that corresponds perfectly with an additive model. Among the remaining subjects, the most common trade-off made was between probability and effort depending on the level of consequences.

\(^{6}\) The effort factor was explicitly included in the CJM task, but not in the SAC task. The main reason for including effort was to explore what kinds of trade-offs might be made between current and future costs of auditing.

\(^{7}\) Approximately five percent of the subjects indicated by their rankings that both probability and consequence were equally important and three percent indicated no clear preference among the three factors.
A second task, the Specific Audit Case (SAC) task, is similar to the research approach taken by others who have studied auditors' planning judgments in an experimental setting. Subjects were given case materials that described a hypothetical client and provided accounting information necessary for audit planning decisions. This task addressed all three potential explanations described above. First, this task provided an additional examination of the auditor's consideration of both probability and consequence, but in a client specific setting. Second, the subject's type, as determined by his or her CJM rankings, was included as an independent variable in a model that attempted to explain variation in audit planning judgments. Third, alternative dependent measures were included to determine if probability of error was able to explain relative concern levels and relative substantive testing levels better than it explained audit hours.

The results of the SAC task seem to most strongly support the third explanation offered. For the accounts receivable account, probability was able to explain alternative dependent variables, but not audit hours. Consequence level provided significant explanation for some accounts receivable planning judgments. However, probability and consequence were highly correlated so consequence did not generally provide significant additional explanation after probability was considered. Results related to a second account, property, plant and equipment, indicated that the planning judgments differ across accounts.

1.3 Expected Contributions of this Research

This research is expected to contribute to the existing auditing research in several ways. In addition to potentially providing insight into the relationship between probability of error and the planned number of audit hours for an account, this research is expected to contribute as described in the subsections below.
1.3.1 Provides Link between Theoretical and Experimental Audit Planning Research

The theoretical auditing research indicates that the consequences associated with not finding an existing error should positively influence the extent of substantive testing planned for an account. The previous experimental audit planning research has not explicitly included consequence as a variable in attempting to explain planned audit hours.

The research described in this dissertation adds consequence as a variable that is assessed by participating auditors and included in a regression model that attempts to explain the audit hours planned for an account. This additional variable is included to discover whether the theoretical result is supported by audit planning judgments made in an experimental setting.

1.3.2 Provides Experimental Evidence regarding Substantive Testing across Accounts

The previous experimental audit planning research has asked subjects to provide judgments for only one account, either accounts receivable (AR) or inventory. Both of these accounts are very important in the audit planning process. Both of these accounts may be assessed as "high consequence" accounts. It may be that the auditor uses different criteria for accounts that are considered less important. To explore the potential differences in judgment processes for other types of accounts, a second account, property, plant and equipment (PPE), is included in the SAC task. During discussions with consulting auditors for this study, most said that PPE is typically not considered to be an important account because there are not many transactions, PPE is not generally included in debt covenant provisions, and the account is not highly susceptible to fraud. By including PPE in addition to AR the results should help to increase our knowledge of the audit planning process in general rather than just increasing our knowledge of the planning process for AR.
1.3.3 Provides Evidence about Auditors' Priorities in Planning an Audit

The CJM task, although a somewhat sterile design, should provide some evidence about what factor the auditors consider to be the most important in planning substantive testing levels for accounts. The auditors' rankings of the 27 combinations will also provide evidence about the commonality of priorities. Analysis of the rankings will indicate whether auditors might consider making trade-offs among probability, consequence and effort. Although these results cannot be used to predict what an auditor would do in an actual audit, it should provide useful information to audit firms about the stated priorities of auditors. Audit firms may be interested in the degree of consensus concerning whether probability or consequence is most important and whether there is a potential for trade-offs to be made in an audit.

1.4 Organization of the Dissertation

This chapter provides background information for audit planning and an introduction to the present research. The second chapter reviews the related literature, focusing on theoretical and empirical approaches to studying audit planning judgments. In the third chapter the research hypotheses are developed. The fourth chapter outlines the research approach. The fifth chapter describes how the research instruments were developed and administered for both tasks. Results are presented in the sixth chapter. The seventh and final chapter includes a discussion of the results as well as suggestions for future research.
The motivation for this dissertation research derives from a sense of incongruity when considering various approaches to audit planning research. Both archival and experimental approaches emphasize the importance of the probability of error in accounts. However, a strong link has not been witnessed between the probability of material error and the number of audit hours planned for an account. As described in section 1.2, a potential explanation for the elusive significant association may be that there are one or more other factors the auditor considers when making audit planning decisions, that have not been explicitly considered in previous attempts to study these judgments empirically.

Decision-theoretic and game theory approaches to studying auditors' planning decisions suggest that "consequences" may be an important factor considered by auditors. In addition to the probability of an error, decision-theoretic and game theory models include the probability of and/or the cost of the auditor making a type II error by accepting an account balance that is materially misstated. Litigation literature and archival error research suggest that the cost of a type II error or level of consequences may be different for different accounts. A combined view of the theoretic literature and the litigation and archival error literature seem to support the premise that the difference between substantive testing levels of accounts within an audit may be related to the perceived consequences associated with not detecting a material error in an account as well as the probability of a material error.
The next four subsections of this paper review the previous accounting literature in the following four areas: 1) previous experimental and archival research in audit planning, 2) the theoretical research on auditor decision-making as it relates to audit planning, 3) archival research on errors, irregularities and litigation and 4) previous research on the auditor's consideration of consequences.

2.1 Audit Planning Research

2.1.1 Experimental Audit Planning Research

Several studies address the auditor's assessment of the likelihood of material error in an account and/or the number of audit hours he or she would plan for an account. The results of these studies indicate much variability in both likelihood assessments and audit hours and a much weaker association between the two measures than is expected.

Kaplan and Reckers' (1984) research focuses on the effect of information received at two different stages of the audit planning process on the auditor's assessment of the likelihood of material error in a hypothetical client's accounts receivable system. First, the subjects received a case describing a client. Within the case were two pieces of information that were varied between subjects. Auditors of two different ranks, seniors and managers were randomly assigned to each of the four combinations of the two manipulated variables. Therefore, the design was a $2 \times 2 \times 2$ with high and low levels of management integrity, high and low control consciousness and senior and manager levels of auditor subjects. The subjects' previous experience with errors was included as a covariate. The subjects were asked "How likely do you believe it is, ex-ante, that the accounts receivable system will lead to a material error?" The results indicated that control consciousness was significant for seniors, but not for managers. The management integrity variable was not significant. The auditor's experienced level of errors (general practice priors, GPP) was significant.
Subjects then received more information about the accounts receivable system and a completed internal control checklist for accounts receivable. The subjects were asked to again estimate the likelihood of material error in the accounts receivable system. The assessments did not change much from the first round. Again, only the GPP variable was significant for both seniors and managers.

This pattern of results implies that the most significant influence on the auditors' initial assessment of the likelihood of error is their personal experience. There are a few reasons why this result may have occurred here in particular. First, the client in the case was described as a first time audit. When a first audit is approached, without the benefit of a history with the client, it may be appropriate for the auditor to rely on his or her previous experience with other clients. A second reason why general practice priors were the dominant influence on the auditors' assessments could be a function of the case description. If portions of the descriptions that were intended to influence the auditors' judgments were not perceived as being realistic, auditors may have ignored them and fell back on their own experience to make a judgment in this case. The significance of GPP may be an artifact of the experiment and not be as significant in practice as it appeared to be in this study.

Kaplan and Reckers (1984) found much diversity in auditors' assessments of the likelihood of error given the same situation. Coefficients of variation ranged from 78 percent to 97 percent for each level of the three independent variables. My research addresses one of the suggestions made by Kaplan and Reckers after commenting on the diverse assessments. They suggest that "...serious attention seems appropriate to efforts to partition subjects into groups that may capture factors and covariates coincident with factor manipulation sensitivity." In an attempt to do this, my study includes a type variable, which indicates whether the auditor believes probability or consequence to be the most important factor when planning substantive testing levels.
Kaplan and Reckers (1989) studied the information search tendencies of auditors, with varying experience levels, in the initial planning stage of the audit. Subjects were given background information about a company and three ratios (gross margin, current ratio and quick ratio) for both the prior year and the current year (unaudited). The ratios for the current year included a seeded error of an unrecorded purchase of inventory. Subjects were asked to first indicate the relative likelihood that the pattern of ratios was the result of an error or irregularity, or a normal year-to-year variation in accounts. A second question asked that the subjects choose ten questions from a list of twenty (ten were related to errors and irregularities and ten were related to year-to-year variations that could have occurred). The ten chosen questions were supposed to represent the questions to which the auditors would seek answers first. From the list of ten chosen questions, the auditors were then asked to rank from one to six, the first six questions they would pursue.

The results showed that less experienced auditors thought it was more likely that the pattern of ratios was the result of an error or irregularity. The more experienced auditors assessed a greater likelihood for year-to-year variations and a lesser likelihood of an error or irregularity. In response to the second requirement of the task, auditors, on average, chose approximately five error questions and five year-to-year variation questions. The third requirement of the task, the preference ranking from one to six of questions chosen, indicated that the less experienced auditors' rankings were consistent with their initial belief about the cause of the unexpected fluctuation in the ratios. Even though more experienced auditors thought the explanation for the ratio pattern was most likely to be normal year-to-year variations, they also ranked error questions among their top six choices of questions to pursue.
Kaplan and Reckers (1989) explained these results in a cognitive framework. They suggested that less experienced auditors followed a confirmation strategy; searching for answers to questions that were consistent with their initial hypotheses (greater likelihood of error or irregularity). More experienced auditors, on the other hand, knew that year-to-year variations (or economic explanations) are more common than errors or irregularities, but still intended to pursue answers to both types of questions in equal numbers. Kaplan and Reckers (1989) attributed the investigation strategy of the more experienced auditors to their more developed cognitive structure of the audit environment. More experienced auditors pursued a more conservative strategy, claimed Kaplan and Reckers. The definition of conservatism invoked by Kaplan and Reckers was one provided by Gibbins (1984, p. 109)

"Because they are generated by a structure which accumulates past experience, response preferences will display a conservative tendency, that is, they will tend to be more stable than is the environment."

An alternative explanation for the more experienced auditors not choosing questions in a manner consistent with their likelihood assessments is that they were considering the consequences of their strategies. An error or irregularity may be perceived as having a higher consequence if not detected than a year-to-year variation. This notion is supported by Palmrose’s (1987) results, as discussed in section three of this chapter. Even though the probability of an error or irregularity was perceived to be fairly low by the more experienced auditors, the high consequence level may be what influenced their information search strategy. My hypotheses about auditor behavior in the audit planning stage would provide an economic explanation for Kaplan and Reckers' (1989) results, rather than the cognitive explanation they provide.

Cohen and Kida (1989) had 96 subjects (50 senior auditors and 46 managers) participate in an experiment on the effect of analytical review results and reliability of
internal control on the auditors' estimates of audit hours for the sales and collection cycle of a wholesale company. The experiment was a 2 X 2 X 2 between subjects design. The internal control system was described as either weak or strong. The analytical review results either indicated that there was an error or that there was not an error. The auditor's experience level (senior or manager) was the third independent variable.

Consistent with the finding of Kaplan and Reckers (1984), Cohen and Kida (1989) found that senior auditors were more influenced by the quality of internal control than the managers were. Managers were influenced more by the analytical review results. A number of audit hours, given to subjects as a benchmark, was described as being what would be appropriate for a similar company that had adequate internal controls, for which analytical review procedures had not yet been performed. Seniors and managers both adjusted audit hours upward from the benchmark as a result of poor internal control or analytical review results indicating an error. Most auditors did not, however, adjust hours downward from the benchmark when internal control was good and/or the analytical review results indicated no error. Cohen and Kida (1989) suggested that auditors may exhibit a conservatism bias; they do not tend to decrease audit hours for an account below some minimum level. The benchmark number of hours may have represented to them that minimum level. This explanation would be consistent with the idea that even though the perceived probability of error may have decreased, there may still have been a high enough perceived level of consequences that influenced auditors to maintain at least the benchmark number of audit hours.

Alternatively, the perceived reliability of analytical review results could have been a factor. If analytical review procedures were not considered to be highly diagnostic, it may be that the auditor was justified in not reducing the level of substantive testing when the analytical review results appeared to indicate no error. The lack of an
indication of error in analytical review results may not be enough of a defense or justification for decreasing testing. This conjecture supposes that the auditor considers the likelihood of making a type II error and does not believe that non-error analytical results reduce this risk enough to warrant reduced testing. The auditor may be considering the potential increase in expected cost of making a type II error against the savings that might be achieved through a lower level of substantive testing and not consider the potential savings in current costs to be worthwhile. Since Cohen and Kida (1989) did not ask for the auditors' assessment of the probability of error, we do not know what impact the internal control and analytical review results had on that assessment or what effect the assessment may have had on the audit hours planned. Therefore, support can not be found to distinguish between the two alternative explanations presented above.

Bedard and Wright (1992) studied the effect of prior audit adjustments and current risk factors on the auditor's assessment of the likelihood of material error and on the level and quality of planned audit hours for an inventory account. Bedard and Wright provided auditor subjects with case materials that described a hypothetical medium sized manufacturing client. Within the case, three independent variables were manipulated in a 2 X 2 X 2 design. The three variables were the presence or absence of a prior year adjusting entry, the presence or absence of management earnings targets and a complex or not complex cost accounting system. The percentage of the auditor's time spent recently on clients with inventory was also included to capture the auditor's level of specific experience. Three dependent variables were measured. Auditors were asked to assess the likelihood that a material error exists given the client information they received. The second dependent variable was the total number of audit hours the auditor indicated he or she would plan for the inventory account for the client. The third dependent variable was a measure of the quality of the audit tests planned. The
effectiveness of a test was determined by its ability to detect the error that the authors seeded into the inventory account.

The first regression model in Bedard and Wright (1992) used the assessed likelihood of material error as the dependent variable. The main effects of the prior period adjusting entry and the auditor's specific experience level were significant, but so were two interactions involving these variables. The specific experience by prior period adjusting entry interaction variable was statistically significant, implying that the presence of a prior period adjusting entry reduced the positive relationship between the auditor's specific experience level and the assessed likelihood of material error. A significant interaction between specific experience and management earnings targets implied that when targets were present, the more experienced auditors tended to assess the likelihood of material error as higher than when the target was not present. The combined results suggested that the more experienced auditors focused on the risk factor of the earnings targets and the less experienced auditors focused on the existence of a prior period adjustment when assessing the likelihood of a material error in the account.

The regression models that used the total number of audit hours and the quality of the planned audit tests as dependent variables produced similar results to each other. The independent variables were the prior period adjusting entry, management earnings targets, the complexity of the cost accounting system, the specific experience level of the auditor, the auditor's assessed likelihood of material error and interaction terms involving specific experience and prior period adjustments, specific experience and earnings targets, and specific experience and the complexity of the cost accounting system. None of the independent variables was able to significantly explain the variation in planned audit hours. Even the auditor's own assessment of the likelihood of material error was not able to provide any explanation of the amount of testing that would be planned. Auditors with more experience in auditing inventories were better able to choose tests that would have
uncovered the error. It is not clear that the auditors believed that there was a high probability of the overstatement error for which they were planning tests. The situation could be similar to that in Kaplan and Reckers (1989) where even though more experienced auditors did not think that the likelihood of such an error or irregularity was very high, it would cost too much (the expected consequences would be high) if they ignored the possibility.

The results of the audit planning studies in an experimental setting seem to imply that the probability of error, as indicated by current risk factors, may not be the major factor considered by auditors when planning the audit of an account. It is hoped that the current research will help to shed light on the results of the previous work discussed above.

2.1.2 Archival Audit Planning Research

Bedard (1989) surveyed audit seniors from three of the Big Eight firms regarding the audits of 48 clients in the retail, wholesale and manufacturing industries. Her questionnaire asked for information on the accounts receivable, inventory and accounts payable accounts. Bedard asked participants to provide the final list of substantive tests performed from the previous year's audit of the client, the initial list of substantive tests planned for the most recent audit and the final list of substantive tests for the most recent audit. Participants were also asked to indicate the nature of and reason for any changes made between the final list of the previous audit and the initial list of the most recent audit and between the initial and final lists of the most recent audit.

Over the two years, there were 38 changes made in substantive testing for the surveyed clients in a total of 142 accounts. Most of these were made between years. There were only six changes made between the initial plan and the final plan for the most recent audit. There were about equal numbers of increases and decreases in testing. The
most common reason given for increasing testing was analytical review results. The most common reason given for decreasing testing was that past audit results were favorable. As Bedard points out, the reason given for decreasing audit tests is potentially troublesome. It implies that the auditor decreases the intensity of the audit when previous audits have not found many errors. If this is true and if the auditee is aware of this strategy, he or she may respond by increasing the incidence of irregularities.

Mock and Wright (1993) surveyed in-charge auditors for 159 audits of manufacturing and merchandising firms of one public accounting firm over a two year period. They measured both changes in the nature of audit testing from one year to the next and changes in the extent of audit testing from one year to the next. They attempted to associate the level and changes in testing with levels and changes in risk factors of the clients. Both the accounts receivable and accounts payable accounts were studied.

The results indicate little variation in the nature of the tests across clients and across periods for the same client. The level of audit testing seemed to be related to the level of inherent risk within the accounts. A highly significant explanatory factor for the level of audit testing planned was the existence of prior errors in the account. Changes in the significant inherent risk factors, however, were not associated with changes in the level of testing. The most common explanation provided by auditors for making changes to the audit plan was to increase the efficiency of the audit.

Overall, the results of the Mock and Wright (1993) study are consistent with those of Bedard (1989). Both indicate that audit plans for a particular client change little from one period to the next. This result seems to underscore the importance of learning more about how auditors initially determine the extent of auditing for each account.
2.2 Theoretical Research

The two most recent streams of theoretical research related to planning the audit are the decision-theory and game theory approaches. Although the decision-theoretic approach includes similar parameters to the later game theory formulations, it does not take into account the auditee's incentives. The decision-theoretic approach takes the probability of error as exogenous to the model, whereas the game theory formulation derives the probability of error in equilibrium as the result of the auditee reacting to the expected actions of the auditor, who anticipates the auditee's reaction. In both approaches, the probability of error, the cost of audit testing and the consequences associated with not finding an existing error are components of the planning decision models.

2.2.1 Decision-Theoretic Research

Kinney (1975) presents a decision-theoretic formulation of the dollar value sample size decision; a decision that determines, in part, the extent of substantive testing. His model includes the a priori beliefs about the existence of error in an account, the costs of auditing the account, the client's reported balance and the costs of a type II and type I error. First, the auditor decides whether to accept or reject the account balance based on his or her prior knowledge or to sample the account. If the auditor chooses to sample the account, the second stage decision is whether to accept or reject the account balance based on the results of the sample. Kinney (1975) assumes a risk-neutral auditor; therefore, it is predicted that the auditor will maximize his or her utility by choosing the alternative with the lowest expected cost.

Kinney (1975) chose various combinations of parameter values to illustrate how his model compared to previous confidence interval approaches for determining sample
sizes. His contribution is in explicitly including the probability that there is no error and both the costs of sampling and the costs of making type I and type II errors. The results indicate that when using his approach, the optimal sample size is more sensitive to estimation errors in the cost of sampling than to estimation errors in type I and type II error costs. Sampling costs are the easiest of the costs to estimate. Since the decision is less sensitive to the accuracy of type I and type II error costs, the inability to accurately predict these costs should not prohibit their inclusion.

Menzefricke (1984) shows, in general, how the optimal sample size can be determined by considering the cost of sampling and the cost differential between a type II and type I error. His formulation is very similar to that of Kinney (1975). In addition, Menzefricke demonstrates the solution to the optimal sample size problem for three probability distributions of errors in an account and two different loss functions. The results of the six combinations of distribution and loss functions indicate that the optimal sample size depends on both specifications.

Menzefricke's (1984) study implies that it may be important to learn more about the nature of the auditor's loss function because it impacts the optimal sampling decision. The current research is intended as a step in that direction. Although the decision-theoretic research indicates that type II costs are relevant in determining optimal solutions to planning decisions, previous experimental research on audit planning has not explicitly included this variable. My research explicitly includes consequence or type II error to study how important it is to auditors and how it affects their planning decisions.

2.2.2 Game Theory Research

Shibano (1990), in recognizing that the audit risk model as specified does not distinguish between intentional and unintentional misstatements, proposes two separate audit risk models: a strategic audit risk (SAR) model and a non strategic audit risk
(NSAR) model. The SAR model relates to intentional misstatements, also called irregularities. The NSAR model relates to unintentional misstatements, also called errors.

The decision-theoretic approaches to audit planning (i.e., Kinney 1975 and Menzefricke 1984) do not take into account the incentives of the auditee and the reaction of the auditee to his or her beliefs about what the auditor will do. In short, the decision-theoretic approach does not address irregularities because the models do not include the auditee as a strategic player who has incentives to intentionally misstate accounts.

Others have modeled the audit process using two-person game theory with the auditor and auditee as strategic players (see for example, Fellingham and Newman 1985 and Newman and Noel 1989). Shibano's (1990) contribution is that he includes not only a hidden action (moral hazard) possibility, but also a potential for hidden information (adverse selection).

In Shibano's models, an account is first classified as being "error-prone" or "irregularity-prone." For an error-prone account, an auditor would assess NSAR. For an irregularity prone account, the auditor would assess SAR. For both types of audit risk, the control risk assessment is assumed to be the same.

Control risk is modeled as a moral hazard problem using a Bayes Nash formulation with the auditor and the controller as strategic players. The controller affects the distribution of audit evidence by maintaining or not maintaining the internal control system. Shibano assumes that, if maintained properly, the control system will detect all errors and irregularities after they occur and that the auditee's management cannot override the controls. He also assumes that information learned during control testing has no effect on substantive testing. These assumptions create a clear separation between control risk and inherent risk8 for the purposes of Shibano's models.

---

8 Since my research concentrates on substantive testing, the review of Shibano (1990) will focus on inherent risk and detection risk only.
The inherent and detection risk components of NSAR are modeled by decision-theoretic techniques. For NSAR, the auditor can rely on his or her prior beliefs. He or she determines the subjective probability that the reported balance is valid given an estimated probability of error and a known distribution of possible true values of the account balance. The prior beliefs and the audit technology employed, along with the auditor's liability ratio, determine the auditor's optimal decision rule and NSAR. The results indicate that NSAR decreases as the auditor's liability ratio increases.

The inherent and detection risk components of SAR are modeled as tests for hidden information. The true balance in an account is determined by nature and the auditee decides whether or not to report it truthfully. For an irregularity account, if the auditee has a low disincentive ratio (his or her payoffs for misreporting are high), the auditor's decision rule and SAR depend only on the auditor's liability ratio. When the auditee's payoffs from misreporting are not so high, the decision rule and SAR depend on the auditee's disincentive ratio, the auditor liability ratio and the audit technology.

This formal derivation of an expanded audit risk model implies that consequences will have an impact on the auditor's planning decisions. Regardless of the auditor's prior beliefs about the probability that there is an error or irregularity, the auditor's liability ratio has a major influence on the auditor's decision rules. The liability ratio is analogous to the "consequences" factor since it is the ratio of the expected loss from a type II error to the expected loss from a type I error. This finding lends support for the expectation that consequences will be an important factor to subjects in my audit planning exercises.

Matsumura and Tucker (1992) model a sequential game between a client's manager and an auditor. They then test the analytical model in an economic experiment. In their model, they examine the effect of four independent variables on actions taken by
a manager and an auditor in a fraud detection setting. The independent variables are: the auditor's loss of money or reputation, requirements of the auditing standards in place, the quality of the client's internal control structure and the audit fee. Before taking any actions, both the manager and the auditor know that unintentional errors occur according to a normal distribution with a known mean and variance. The manager first chooses whether or not to commit fraud. The auditor then decides whether or not to test transactions (tests of controls) and learn the true rate of unintentional errors. If the auditor does not plan to test controls, he or she then chooses a high or low level of detailed tests of balances based on the a priori distribution of errors. If the auditor decides to test controls, the desired level of detailed tests of balances is based on the true unintentional error rate discovered through the test of controls. According to Matsumura and Tucker's model, the auditor can detect fraud only if he or she performs detailed tests of balances; tests of transactions are assumed not to result in fraud detection. The probability that the manager will commit fraud is based on the manager's expected gains and expected losses associated with committing fraud. The auditor chooses an audit strategy based on his or her prior belief about the probability that the manager will commit fraud. Essentially, if the prior belief is low, the auditor is expected to choose a low level of detailed tests of balances. If the prior belief is high, the auditor is expected to choose a high level of detailed tests of balances. In other cases (moderate prior belief), the auditor is expected to test transactions first and then choose the level of detailed testing based on the updated error information.

For different combinations of auditor penalty, required auditing standards and quality of internal control, Matsumura and Tucker (1992) show the ranges of probability of fraud for which the auditor chooses each strategy. The three strategies are: tests of controls and then detailed testing, only detailed testing at a low level and only detailed testing at a high level. Germane to my research is the change in strategy when the
auditor's penalty increases. Matsumura and Tucker's (1992) results indicate that the probability of material fraud for which the auditor will choose a high level of detailed tests of balances (substantive testing) decreases as the audit penalty increases. More specifically, the auditor increases detailed tests of balances in reaction to a higher auditor's penalty even when the quality of the internal control system remains constant. In table 2 from Matsumura and Tucker (1992, p. 761), a comparison of two cases that have the same internal control characteristics, case 6 (high auditor penalty) and case 8 (low auditor penalty) reveal the impact of the audit penalty. For case 8, a high level of testing would be planned when the probability of fraud is ≥ .81, but for case 6, where the penalty is increased, a high level of testing would be planned when the probability of fraud is ≥ .35.

Undergraduate business students participating in Matsumura and Tucker's economic experiment increased detailed testing of balances as predicted when the auditor's penalty was increased.

Matsumura's and Tucker's (1992) research supports the idea that auditors would increase substantive testing levels if they perceived an increase in expected future loss even if the probability of error remained constant.

2.3 Archival Research on the Incidence and Consequence of Errors and Irregularities

Palmrose (1987) reports on a sample of more than 450 legal cases brought against the fifteen largest auditing firms in the years 1960 - 1985. The data indicate that management fraud was associated with about 44 percent of all litigation against auditors in this period, defalcations were associated with 2 percent and error with 48 percent (6 percent of the cases are not classified). The major finding of Palmrose with regard to resolved cases was that, of those that involved management fraud, most resulted in the
auditor paying a settlement or judgment (74 percent). Of those that involved errors, the majority were dismissed (52 percent). This finding may explain why auditors might consider consequences to be important and why higher consequences might be assessed for accounts that are more likely to be subject to fraud.

Hylas and Ashton (1982) report their findings about audit adjusted errors in 152 audits of one of the Big Eight public accounting firms. The authors asked auditors to provide information on three aspects of these errors. The account or cycle where the error took place, the event that led to the initial discovery of the error and the cause of the error, were recorded by auditors. They were also asked to indicate whether they thought each error was intentional or unintentional.

Of the total errors recorded for the Hylas and Ashton (1982) study (one firm's clients for the year ending December 31, 1978), 16 percent of the errors were in the revenue cycle and accounts receivable and 10 percent were in property, plant and equipment account. Twenty-six percent of the errors were caused by personnel problems (17.4 percent by new or inexperienced employees) and another 15 percent by client personnel having insufficient accounting knowledge. Judgment error by the client was cited as the cause of 15 percent of the errors. Only 10 of 281 errors (3.6 percent), found on 152 audits, were judged by the auditors on the engagements involved to be intentional errors. Very large errors (more than two percent of total assets) were found more often in accounts receivable and property, plant and equipment.

Kreutzfeldt and Wallace (1986) asked auditors to complete a questionnaire regarding the latest audit for 260 clients of Arthur Andersen (one of the Big Eight firms). A total of 1,506 adjusting journal entries and passed adjusting journal entries were studied. In this sample, more errors were found in small companies than in larger ones.
The greatest number of errors were found in receivables (15 percent), inventory (11 percent), accounts payable (9 percent), accrued liabilities (16 percent) and fixed assets and accumulated depreciation (17 percent).

The most prevalent error in accounts receivable was due to judgmental evaluation of bad debts. The most common error in the fixed asset account was classified as a misapplication of generally accepted accounting principles (GAAP) and involved recording an expense as an asset or recording an asset as an expense. Accounts receivable had more overstatement errors (66 percent) than understatements while errors in fixed assets were evenly divided between overstatements and understatements. Consistent with the finding of Hylas and Ashton (1982), about one-third of all errors detected were the result of personnel problems. High turnover of accounting personnel was associated with a significantly higher error rate.

Wright and Ashton (1989) extended the Hylas and Ashton (1982) study and found similar results. Wright and Ashton (1989) included waived and booked errors that were at least twenty percent of gauge ($1.6 \times (\text{the larger of total assets or net revenues})^{2/3}$). The gauge calculation represents the average level of planning materiality for the audit firm that provided the data. Wright and Ashton (1989) report the results of error detection by auditors in 186 engagements with year-ends of September 1984 through March 1985. They report that approximately 9 percent of the errors were found in accounts receivable, approximately 8 percent involved the revenue account and approximately 5 percent were in the property, plant and equipment account.

In contrast to Hylas and Ashton (1982), Wright and Ashton (1989) found a much smaller percentage of the errors caused by personnel problems, approximately 6 percent compared to 26 percent in Hylas and Ashton. Insufficient accounting knowledge
accounted for almost 29 percent, compared to 15 percent in Hylas and Ashton, and judgment accounted for 20 percent of the errors, compared to 15 percent in Hylas and Ashton.

The archival error studies provide support for the research design (described in Chapter V). Each of the accounts used in the SAC task has been found to have high incidence and large dollar amounts of error. The risk factors included in the case are ones that have been found to be frequently associated with material errors in the accounts.

Loebbecke, Eining and Willingham (1989) (hereafter, referred to as LEW) extended an earlier paper by Loebbecke and Willingham (1988) in which they developed a model to detect management fraud. LEW (1989) tested the model by determining whether it would have resulted in the detection of management fraud in engagements of the firm involved in the study. They surveyed auditors to determine what their experiences were with respect to management fraud and how their experiences corresponded with the elements predicted to be present by the model.

The Loebbecke and Willingham (L & W) model gives auditors a way to think about the potential for management fraud. The model indicates that three things are necessary for there to be a non zero probability of a material irregularity. The person committing the fraud must have the opportunity to do so, a motivation for committing fraud and an attitude that would allow him or her to commit fraud. If all three of these elements are present, the auditor should assess a high probability of management fraud. If any one of the three elements is missing, the probability of management fraud should be assessed at zero.

LEW presented 31 risk factors derived from SAS No. 53 (The Auditor's Responsibility to Detect and Report Errors and Irregularities) that are intended to signal the potential for the existence of a misstatement (error or irregularity). For each of the risk factors, the authors indicated whether any of the three elements required for
suspecting an irregularity were satisfied. LEW, therefore, showed how the list of risk factors provided by SAS 53 could be further analyzed to determine which ones might indicate the existence of an error (unintentional misstatement) and which ones the existence of an irregularity (an intentional misstatement).

The survey resulted in responses from 121 audit partners who had some experience with irregularities. From information provided by the partners about irregularities they had encountered, LEW described the nature of surveyed irregularities and tested the L & W assessment model against the irregularities found. LEW showed that application of the L & W assessment model would have been successful in indicating fraud in 78 percent of the 77 material management fraud cases.

Data about the irregularities indicated that almost 20 percent of all management fraud instances encountered were in the manufacturing industry. The most common action involved in the instance of the fraud was the overstatement or incorrect valuation of assets (57.7 percent of management fraud events reported). About 20 percent of the fraud cases involved the revenue cycle, about 14 percent involved the accounts receivable account and about 6.5 percent involved the property, plant, and equipment account. Substantive tests of detail were successful in detecting almost 56 percent of the instances of management fraud.

One fourth of the management fraud cases for which the auditors provided detailed information (88 fraud cases in total were detailed) involved "accounts that are material to the financial statements for which extensive judgment is involved in determining their balances." Twenty-six of the detailed cases involved companies in a period of rapid growth. These two characteristics are examples of those that provide opportunity for the fraud to take place. Motivational factors include "compensation arrangements based on performance" (9 out of 88) and "significant contractual commitments to which the company is subject" (18 out of 88 cases). Two of the attitudes
were "management displays an overly aggressive attitude toward financial reporting" (29 out of 88 cases) and "management has engaged in frequent disputes with the auditors" (14 out of 88 cases). These statistics indicate that it would not be unreasonable to suspect fraud in the audit scenario used in the SAC task, as described in Chapters V.

2.4 Auditors' Assessments of Consequences

Two studies (Ward 1976 and Lewis 1980) addressed the preferences of auditors with respect to outcomes of auditing decisions, but were not directly related to planning the audit. Ward asked 24 auditors (partners and managers) to order twenty potential outcomes of an error not being detected by placing each item into one of five groups, from most severe to least severe consequence.

The average response of participants indicated that the "degree to which audit complied with professional standards" was the most important outcome. The ratings by participants in this study indicated high consensus among the auditors. When given the potential outcomes, the 24 partners and managers generally agreed on the relative importance of the outcomes. This indicates that these auditors have similar value systems. The problem is that within an audit, the auditor is not provided with these outcomes, but with audit situations that could result in one or more of these outcomes. It is not clear how well auditors can predict the outcome of not detecting an error in a particular situation or whether the auditors' assessments of the audit situations would result in consensus outcome predictions.

Lewis (1980) asked auditors for their preferences over certain action/outcome pairs in a disclosure setting where there was a product liability suit filed against the client. Lewis asked 76 auditors to rate the desirability of six possible outcomes. The six outcomes were formed by three actions the auditor could take and two results of the suit.
The auditor could choose to accrue the liability, disclose the lawsuit in a footnote or not disclose the lawsuit at all. The lawsuit could result in the client being held liable or the client not being found liable. Most of the auditors ranked the outcome of "no disclosure, client liable" as the least desirable combination. This study indicates what the auditor might do based on his or her assessment of the probability that the client would be held liable, but we do not know what type of consequence the auditor would predict to occur as a result of that outcome.

From Ward's (1976) study, we learn how auditors value the potential effects of a non-detected error. From Lewis' (1980) study, we learn which action/state combination is least desirable. The auditors in Lewis' study would least prefer to make a type II error. My study asks auditors to assess the level of consequences associated with making a type II error in a particular audit situation. My analysis of the subjects' responses includes exploring the relationship between the perceived level of consequence and the auditor's planned level of substantive testing which would be aimed at reducing the possibility of making a type II error. The design of my research does not allow one to determine the specific consequences an auditor expects, but only how severe the consequences are expected to be. The results will show only how important consequences are, compared to the probability of error in an account.
2.5 Summary of Relevant Literature

The empirical audit planning research has found neither consistent nor strong associations between the probability of error and the number of audit hours planned for an account. The decision-theoretic and game theory approaches to audit planning indicate that the level of consequences (or the cost of making a type II error) should be an important influence on planning levels. Litigation research and error and fraud research suggest that the level of consequence may differ by account. Another stream of research suggests that auditors consider type II errors to be a very important influence on their audit decisions. The literature as a whole supports the research hypotheses described in Chapter III.
CHAPTER III
DEVELOPMENT OF RESEARCH HYPOTHESES

The research hypotheses addressed in the dissertation relate to the three potential explanations for the lack of significant relationship between probability of error in an account and the number of audit hours planned for the account. Section 3.1 relates to the possibility that the auditor considers consequences, in addition to probability of error, when making planning judgments. Section 3.2 addresses the potential effect of a particular individual difference on the resulting planning judgments. Section 3.3 addresses the potential impact of using the number of audit hours as the dependent variable.

3.1 Effect of Probability and Consequence Levels on Planning Judgments

As explained in Chapter I, one potential explanation for the lack of significant association between probability of error in an account and the audit hours planned for the account is that the auditor also considers the consequences associated with not detecting an error in the account. In this section, a decision-theoretic analysis and subsequent numerical example are used to derive the research hypotheses related to the probability and consequence factors and the interaction between probability and consequence.

3.1.1 Decision Theoretic Model

A decision-theoretic analysis related to the desired level of substantive testing is formulated here in a similar way to Kinney's (1975) and Menzefricke's (1984) analyses of
the optimal sample size problem. Although the sample size decision is a more specific problem, the essence of both problems is the same; the optimal decision is one that minimizes the total cost of auditing including that of making a type II error.

3.1.1.1 Assumptions

In modeling the substantive testing level decision, the following assumptions are made:

1) The auditor is risk neutral\(^9\) and therefore wishes to maximize expected value, or in this case, minimize expected cost.

2) The internal control system will not be relied on, so control risk is set at 1.

3) Detection risk\(^10\) (DR) for a high (H) level of substantive testing is less than detection risk for a low (L) level of substantive testing. (DR\(_H\) < DR\(_L\))

4) The audit effort (E) that would be required for a high level of substantive testing is greater (more audit hours and/or higher level(s) of personnel required) than the audit effort required for a low level of substantive testing. (E\(_H\) > E\(_L\)).

5) The expected level of consequence (C) is the same for not detecting a material error in a particular account whether a high or low level of substantive testing is planned.

6) There are two states of nature possible. In one state, there is a material error and in the other state, there is no material error. The auditor's ex-ante estimate of the probability of a material error (P) is an estimate of the probability that the true state is "material error."

---

\(^9\) This is the same assumption made by Kinney (1975). Menzefricke demonstrated that the optimal sample size was sensitive to the form of the auditor's loss function. Since there is no theoretical basis for choosing a specific utility function, a linear function is assumed here for the sake of simplicity.

\(^10\) Detection risk is defined in Chapter I as the probability that the audit tests will not result in revealing an existing material error.
3.1.1.2 The Auditor's Choice as a Minimization of Expected Cost

The following matrix shows the cost of each state/action combination.

<table>
<thead>
<tr>
<th>Auditor's Action</th>
<th>True State</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Material Error</td>
</tr>
<tr>
<td>High level of substantive testing</td>
<td>$E_H + DR_HC$</td>
</tr>
<tr>
<td>Low level of substantive testing</td>
<td>$E_L + DR_LC$</td>
</tr>
</tbody>
</table>

**Figure 2. Cost Matrix for Level of Substantive Testing Choice**

In order for the auditor to make an optimal decision based on expected costs, he or she would need an estimate of $E_H$, $E_L$, $C$, $DR_H$, $DR_L$ and $P$. The auditor should choose a high or low level of substantive testing based on the expected cost of each level.

If a high level of substantive testing is planned, the expected cost, $EC_H$, can be calculated as follows:

$$EC_H = P(E_H + DR_HC) + [(1 - P)(E_H)] = E_H + P(DR_HC)$$  (1)

If a low level of substantive testing is planned, the expected cost, $EC_L$, can be calculated as follows:

$$EC_L = P(E_L + DR_LC) + [(1 - P)(E_L)] = E_L + P(DR_LC)$$  (2)
To minimize the total cost of auditing an account, the auditor should plan a high level of substantive testing if:

\[ E_{CH} \leq E_{CL}, \text{ or if:} \]
\[ E_H + P(DR_H)C \leq E_L + P(DR_L)C, \text{ or if:} \]
\[ \frac{E_H - E_L}{(DR_L - DR_H)} \leq PC \]

This implies that the level of substantive testing would be increased as long as the cost of increasing testing (and reducing the detection risk) is less than or equal to the expected cost of not detecting a material error.

3.1.1.3 Numerical Example

To illustrate the decision-theoretic approach to choosing substantive testing levels, an example is provided below. The numbers were chosen arbitrarily to represent high and low levels of probability and consequence. When considering an actual audit, the expected consequences would be a combination of the auditor's assessment of the probability that there would be a lawsuit if an error were not detected, the probability that the lawsuit would result in the audit firm making a payment, the estimated size of the payment and the expected dollar impact on the audit firm's reputation as a result of litigation. The cost of testing, the probability of error and the detection risk levels would be estimated by the auditor using his or her knowledge of the account and the client.
The amounts used in this example are:

Low $P = .30$

High $P = .80$

Low $C = $300,000

High $C = $800,000

If an auditor assessed an account as having a low level of probability and a low level of consequence, then he or she would be expected to increase testing as long as the cost of decreasing the detection risk were less than $900 per one percent decrease as shown below:

\[
\frac{E_H - E_L}{(DR_L - DR_H)} \leq 0.30 \times 300,000
\]  

\[
E_H - E_L \leq 90,000(DR_L - DR_H)
\]

\[
E_H - E_L \leq 90,000(.01) = 900
\]

Using the same computational procedure as above, table 1 shows the maximum increase in cost that would be acceptable for decreasing the detection risk by one percent for each of the four combinations of assessed probability and consequence.

Table 1. Maximum Acceptable Cost of Decreasing Detection Risk by One Percent

<table>
<thead>
<tr>
<th></th>
<th>High Probability</th>
<th>Low Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Consequence</td>
<td>$6,400</td>
<td>$2,400</td>
</tr>
<tr>
<td>Low Consequence</td>
<td>$2,400</td>
<td>$900</td>
</tr>
</tbody>
</table>
As can be seen in table 1, when the assessed level of probability (consequence) remains constant at a low level and the assessed level of consequences (probability) is increased, the maximum acceptable cost of decreasing the detection risk increases by $1,500. Alternatively, when the assessed level of probability (consequence) remains constant at a high level and the assessed level of consequences (probability) is increased, the maximum cost of decreasing the detection risk increases by $4,000. This analysis suggests that a change in the level of probability or consequence has an effect on the desired level of substantive testing and also that the acceptable increase in the cost of testing depends on the level of probability (consequence) when the level of consequence (probability) changes.

3.1.2 Statement of Research Hypotheses

The decision-theoretic analysis and more specifically, the numerical example in the preceding section provide support for the following research hypotheses.

**H1:** Ceteris paribus, the level of concern, level of substantive testing and number of planned audit hours for an account will be positively associated with the assessed probability of error for the account.

**H2:** Ceteris paribus, the level of concern, level of substantive testing and number of planned audit hours for an account will be positively associated with the assessed consequence level for the account.
H3: Ceteris paribus, the effect of a change in the level of consequence (probability) on the level of concern, level of substantive testing and number of planned audit hours for an account will be greater when the assessed level of probability (consequence) is high than when the assessed level of probability (consequence) is low.

3.2 Effect of the Auditor's Type on Substantive Testing Levels

The second potential explanation, described in Chapter I, was individual differences between auditors. One such difference was explored in this research; the subject's most important factor. Each subject was classified as a probability or consequence type, or as neither type. A person's type was determined by the judgments made in the CJM task (described in detail in Chapter V). If a subject placed more emphasis on the level of probability (consequence) when ordering the 27 combinations of probability, effort and consequence levels, then he or she was classified as a probability (consequence) type. If a subject placed about equal emphasis on each of the two factors or placed the most emphasis on the effort factor, then he or she was classified as neither a probability (TYPE P) nor a consequence type (TYPEC).

3.2.1 CJM Scaling and the Effect of Type

Part of the analysis of the 27 rankings provided by subjects in the CJM task consisted of assigning scale values to each level of each factor. The scale values are determined so that when they are added together for a particular combination, the resulting total score places the combination in the preference order indicated by the subject. The total score values for combinations are normalized on a 0 to 100 scale. The most preferred combination receives a total score of 100 and the least preferred
combination receives a total score of 0. For example, if someone preferred to plan the highest level of substantive testing for an account with high probability, low effort and high consequence, the values assigned to each of those levels of factors would add up to 100 for that combination.

If a person ranks the combinations in such a way that he or she appears to combine factors additively and place the most emphasis on probability (consequence) and the next most emphasis on consequence (probability), then he or she will rank the combinations according to a PCE (CPE) prototype. Prototype rankings are illustrated in table 2 on the next page.
Table 2. Prototype Rankings

<table>
<thead>
<tr>
<th>CPE Prototype</th>
<th>High Conseq.</th>
<th>Medium Conseq.</th>
<th>Low Conseq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROBABILITY - HIGH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Effort</td>
<td>1</td>
<td>10</td>
<td>19</td>
</tr>
<tr>
<td>Medium Effort</td>
<td>2</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>High Effort</td>
<td>3</td>
<td>12</td>
<td>21</td>
</tr>
<tr>
<td>PROBABILITY - MEDIUM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Effort</td>
<td>4</td>
<td>13</td>
<td>22</td>
</tr>
<tr>
<td>Medium Effort</td>
<td>5</td>
<td>14</td>
<td>23</td>
</tr>
<tr>
<td>High Effort</td>
<td>6</td>
<td>15</td>
<td>24</td>
</tr>
<tr>
<td>PROBABILITY - LOW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Effort</td>
<td>7</td>
<td>16</td>
<td>25</td>
</tr>
<tr>
<td>Medium Effort</td>
<td>8</td>
<td>17</td>
<td>26</td>
</tr>
<tr>
<td>High Effort</td>
<td>9</td>
<td>18</td>
<td>27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PROBABILITY - HIGH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Effort</td>
<td>1</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Medium Effort</td>
<td>2</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>High Effort</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>PROBABILITY - MEDIUM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Effort</td>
<td>10</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>Medium Effort</td>
<td>11</td>
<td>14</td>
<td>17</td>
</tr>
<tr>
<td>High Effort</td>
<td>12</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>PROBABILITY - LOW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Effort</td>
<td>19</td>
<td>22</td>
<td>25</td>
</tr>
<tr>
<td>Medium Effort</td>
<td>20</td>
<td>23</td>
<td>26</td>
</tr>
<tr>
<td>High Effort</td>
<td>21</td>
<td>24</td>
<td>27</td>
</tr>
</tbody>
</table>
The scaling\textsuperscript{11} that results for a subject whose CJM rankings perfectly fit one of these prototypes can be used to illustrate the potential effect of the type variable on planning judgments. The dependent variable used in the CJM task was the relative preference for a high level of substantive testing. The value for the dependent variable is shown in table 3 as a normalized score on a 0 to 100 scale.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
\textbf{Combination} & \textbf{Value of Dependent Variable for:} & \\
 & \textbf{PCE Prototype} & \textbf{CPE Prototype} \\
\hline
Low P, Medium E, Low C & 3.8 & 3.8 \\
\hline
Low P, Medium E, High C & 26.9 & 73.1 \\
\hline
High P, Medium E, Low C & 73.1 & 26.9 \\
\hline
High P, Medium E, High C & 96.2 & 96.2 \\
\hline
\end{tabular}
\caption{Score Values for Combinations of Factors}
\end{table}

An observation from table 3 is that a significant interaction between Type and factor would be expected. For subjects who indicate that probability (consequence) is the most important factor, an increase in probability (consequence) appears to have a larger impact than an increase in consequence (probability). When probability (consequence) increases from low to high for a given level of consequence (probability), the change in the dependent variable for the PCE (CPE) prototype is 69.3 points. But, when consequence (probability) increases from low to high for a given level of probability (consequence), the change in the dependent variable for the PCE (CPE) prototype is only 23.1 points.

\textsuperscript{11} The process of numerical scaling is described in section 5.2.4.
If subjects make decisions in the SAC task that are qualitatively similar to the rankings in the CJM task then it is expected that those who are classified as probability (consequence) types would increase testing more for an increase in probability (consequence) than they would for an increase in consequence (probability). The following hypotheses address the expected effect of this individual difference variable on audit planning judgments.

**H4:** If an auditor is classified as a probability type, his or her assessment of probability of error in an account will have a larger positive influence on the level of concern, relative level of substantive testing and number of audit hours than the probability assessment of someone who is not classified as a probability type.

**H5:** If an auditor is classified as a consequence type, his or her assessment of the consequence associated with not detecting an error in an account will have a larger positive influence on the level of concern, relative level of substantive testing and number of audit hours than the consequence assessment of someone who is not classified as a consequence type.

### 3.3 Alternative Dependent Variables

Three dependent variables are being used in this research. Traditionally the dependent variable for research on audit planning has been the number of audit hours planned. This research adds two additional dependent variables in an attempt to provide some insight into the lack of explanatory power that probability of error was seen to have in Bedard and Wright (1992), Kaplan and Reckers (1989) and Mock and Wright (1993).
Audit hours indicated by auditor subjects in previous experimental audit planning research among auditors responding to the same situation did not seem to vary with subjects' perceived probability of error. The substantive test planning process assumed in conducting this study is illustrated in figure 1 (Chapter I). In the second diagram in figure 1, one can see where the response of planned audit hours may have some variation that is not necessarily a direct result of the auditor's assessment of the level of testing required. One can think of the number of audit hours as resulting not only from planned detection risk for the account, but also from the particular tests planned, the level of audit personnel to be involved and the sample size for each test. By asking for a relative level of substantive testing, the dependent variable should better reflect the auditor's planned detection risk and be more closely related to the assessed probability of error. Hypothesis six, H6, addresses the dependent variable issue.

**H6**: Alternative dependent variable measures, such as the relative level of concern or the relative level of substantive testing will result in a clearer and stronger relationship between the probability of error in an account and the level of testing planned than the traditional dependent variable of total audit hours.
CHAPTER IV
RESEARCH APPROACH

The first section of this chapter provides justification for using an experimental approach to the research issues being studied. Section 4.2 explains why CJM is suitable for one of the tasks and section 4.3 discusses the suitability of multiple linear regression for analysis of the SAC task.

4.1 Suitability of the Experimental Approach to Audit Planning Research

The research approach used was experimental, using practicing auditors as subjects. Since the impetus for the current research derived from reviewing the prior experimental audit planning research, much of the approach was consistent with the existing literature.

In addition to providing a link to previous research, the experimental approach has other advantages over archival and theoretical research approaches. If one is interested in the processes underlying audit judgments, it is imperative to study auditors making judgments. This may be done by looking at judgments that auditors have made on actual audits (archival research) or by creating an audit situation and asking auditors to provide responses that can be studied (experimental research). An experimental approach, rather than an archival approach, allows the researcher to manipulate the levels of the variables and to design a systematic study of the effect of the variables. It allows the researcher to study the judgment processes of auditors from several firms, using a single audit situation. What is gained in controllability however, is lost in generalizability. Archival
research is more generalizable if common relationships are found across several audit situations. An experimental approach presents each auditor subject with the same client and can associate auditors' different reactions to controlled changes in the situation. Subjects from several auditing firms were involved in this study to increase the potential for generalizing the results.

In the scheme of accounting research, the objectives of this research place it under the heading of "policy capturing studies." The basic problem being studied in policy capturing research is how changes in the independent variables (factors) affect the value of the dependent variable (level of substantive testing). Ashton (1982) identifies the three predominant approaches for policy capturing: multiple regression (MR) analysis, analysis-of-variance (ANOVA) models and conjoint measurement (CJM) analysis. This dissertation reports on two tasks and a questionnaire that were administered to auditors from six different firms. One task, the CJM task, was intended to be exploratory in nature. Since CJM analysis requires very few assumptions about the data, it provided the basis for designing and analyzing the CJM task as described below. The other task, the Specific Audit Scenario (SAC) task, was intended to be consistent with the previous literature, which typically has used MR analysis.

4.2 Suitability of Conjoint Measurement Approach

One objective of this research is to learn the relative importance of three factors (probability, effort and consequence) and how their effects are combined to form the basis for audit planning judgments. The preference ordering (or CJM) task was intended to achieve this objective. There was no a priori justification for assuming a particular functional form for the relationship between the variables. It was expected only that auditors would be able to make ordinal preference judgments by account. For these
reasons, CJM techniques seemed most appropriate for achieving a subset of the research objectives outlined in Chapter I. As stated by Krantz and Tversky (1971), "the theory of conjoint measurement is concerned with the qualitative laws that govern the composition of several variables." Conjoint measurement requires that a subject rank order all possible combinations of the factors. The subject's rankings are then analyzed to determine the relative importance placed on each factor, any trade-offs made among the factors and the best candidate for modeling how the factors were combined to form the subject's rankings.

The assumptions of CJM as summarized by Person (1979) and listed below were met by the data collected:

1. Responses are the joint effect of two or more factorially varied explanatory variables.
2. Responses are interpreted as ordinal scale data.
3. Factor levels can be rank ordered for each qualitative variable.
4. There is one response per cell (i.e., combination of factor levels).
5. Each cell response is error free.12

The objective of CJM analysis is to determine if the rank ordered data provided by a subject seem consistent with one of four simple polynomial models: additive, multiplicative, distributive or dual-distributive. A particular model is suggested by the satisfaction of certain axioms.13 Previous research in auditing (Moriarity and Barron 1976 and Schneider 1984) and most of the research in other disciplines have found or assumed additive models (Emery and Barron 1979 and Nygren 1986). The two

12The significance of this assumption and the implications of violations of it are discussed in section 5.2.4.

13 The models are described in Appendix F. The major axioms and a flowchart that relates the axioms and models appear in Appendix E.
applications of CJM in auditing settings are discussed in section 5.2.4.1, "Previous Applications of CJM in Accounting Research."

Since the particular judgments being studied by this research may involve trade-offs being made by the auditors, the interest is not only in finding a paramorphic model but in understanding the trade-offs being made. A computer algorithm called CMSCAL, developed by Nygren (1985, 1986) was used. This algorithm allowed a more meaningful analysis of the rank ordered data than was previously available. Section 5.2.4.2 describes the nature of the analysis available with CMSCAL.

4.3 Suitability of Multiple Regression Analysis of the SAC Task

The design of the SAC Task was similar to that of previous audit planning research. Since the objectives of the current research relate to exploring explanations for the lack of significant results in other research, it was important to have one of the tasks resemble those of the existing research. Consistent with the previous research, the SAC task includes a description of a client and information about the client's accounts receivable account and then asks the auditor subjects to make planning judgments for the client. In addition, the SAC task includes information about a second account (property, plant and equipment) to provide an opportunity to study the effect of different levels of consequences within the same audit scenario.

Similar to the existing research, a multiple linear regression model was used to explore significant associations between the independent and dependent variables. Psychology research has concluded that a simple linear model is usually excellent for paramorphically representing human judgments (Dawes and Corrigan 1974). The objective of this task was to use the same approach as the previous research but to add independent variables and to propose alternative dependent variables to find significant associations that have not been previously documented.
CHAPTER V
DEVELOPMENT AND DESCRIPTION OF RESEARCH

The first section of this chapter describes the overall research design and how the study was administered. Section 5.2 describes the development, administration and analysis of the CJM task. Section 5.3 describes the development and administration of the SAC task.

5.1 Overview of the Study

5.1.1 Consulting Auditors

To assist in the design of the experimental materials, the author consulted with five auditors (from four firms) with experience in audit planning. Four of the auditors were from three of the Big Six auditing firms. The fifth auditor was from a large regional firm. The assumptions made underlying the research and the objectives of the research were discussed. The five consulting auditors provided support for the project by offering their views on audit planning issues, commenting on materials and providing subjects for pilot testing and for the administration of the final version of the experiment.

5.1.2 Administration of the Study and Research Design

The experiment was planned to consist of a 60 to 90 minute session which included the two experimental tasks and a post-experimental questionnaire. Each

14 Although the completion time was not requested from the subjects, a few indicated that it took them about an hour to complete both tasks. There was no feedback indicating that the estimated time was unreasonable.
participant received a large envelope of experimental materials which included an introductory sheet and two manila envelopes. The introductory sheet consisted of general comments about the research (see Appendix D). One of the manila envelopes was identified with a "1" as the envelope that contained materials for the task the participant should do first. The other envelope, marked with a "2," contained the second task. Subjects were randomly selected to do either the CJM task or the SAC task first and the other task second. The envelope containing the CJM task included the instructions for the task and 27 four-inch by six-inch cards. The envelope containing the SAC task included the case materials and questionnaire.

The chart below illustrates the overall experimental design.\textsuperscript{15}

<table>
<thead>
<tr>
<th>Number of Subjects</th>
<th>15</th>
<th>15</th>
<th>19</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Task</td>
<td>CJM</td>
<td>CJM</td>
<td>SAC - Version A High Prob. AR Low Prob. PPE Questionnaire</td>
<td>SAC - Version B Low Prob. AR High Prob. PPE Questionnaire</td>
</tr>
<tr>
<td>2nd Task</td>
<td>SAC - Version A High Prob. AR Low Prob. PPE Questionnaire</td>
<td>SAC - Version B Low Prob. AR High Prob. PPE Questionnaire</td>
<td>CJM</td>
<td>CJM</td>
</tr>
</tbody>
</table>

\textbf{Figure 3. Overall Experimental Design}

\textsuperscript{15} The total number of subjects in all four cells equals 64. A total of 65 auditors participated in this research, but one subject completed only the CJM task and therefore was not included in this table.
5.2 CJM Task

5.2.1 Development of the CJM Task

To administer the CJM task, three levels of each factor were needed. Preliminary discussions with consulting auditors were directed at what they considered when planning substantive testing for the different accounts within an audit. From the items they mentioned and from an audit manual, auditing texts (Defliese, Jaenicke, Sullivan and Gnospelius 1984 and Kell, Boynton and Ziegler 1989) and the auditing standards (AICPA 1989), a description for each level of each factor was developed. Sets of twenty-seven (3 x 3 x 3) four-inch by six-inch cards were made. Each card contained one of the possible combinations of the three factors.\textsuperscript{16} Background information and instructions were written to accompany the cards. The original background information included, for each factor, a description and examples of what one might consider as components of that factor. Two former auditors, who are currently Ph.D. students at The Ohio State University, reviewed the factor level descriptions and accompanying materials for reasonableness. After revisions, the CJM task was pilot tested.

5.2.2 Pilot Testing of the CJM Task

Three of the consulting auditors and several of their colleagues evaluated and performed the CJM task, providing pilot test results for the task.\textsuperscript{17} The auditors' lengths

\textsuperscript{16} The actual descriptions that were used in the pilot test and administration of the experiment were the same and are provided in Appendix A. A sample card appears as figure 7 in Appendix A.

\textsuperscript{17} Thirteen auditors from four auditing firms in Columbus, Ohio completed the CJM task in a pilot test.
of experience ranged from 1 year and 4 months to 20 years. As a result of comments from some of the pilot test participants the instructions for the CJM task were pared down to their current form.\textsuperscript{18} (See Appendix B.)

5.2.3 Administration of the CJM Task\textsuperscript{19}

The subjects were asked to sort 27 cards into the order that represented their preference for planning a high level of substantive testing on the described "accounts." The profile containing level 3 (high probability, low effort and high consequences) on all three factors was pre-ranked as #1 for all subjects. The profile containing level 1 (low probability, high effort, and low consequences) on all three factors was pre-ranked as #27 for all subjects. The profile labeled #1 represents the highest preference scenario; it describes the account on which the auditor would plan the highest level of substantive testing. The profile labeled #27 is the lowest preference scenario and represents the account on which the auditor would plan the lowest level of substantive testing. The other 25 profiles were randomized and put into a stack.

5.2.4 CJM Analysis

There are two facets to conjoint measurement analysis, an axiomatic approach and a numerical approach. The data used as input to CMSCAL (Nygren 1985, 1986), a CJM algorithm, consisted of subjects' rankings of various combinations of factor levels. For example, in the context of this research, one combination consists of level 2 on factor P (medium probability), level 3 on factor E (low effort), and level 1 on factor C (low consequence). Currently available conjoint measurement algorithms can easily handle up

\textsuperscript{18} Reviewers of the dissertation draft proposal also commented that the original instructions provided too much guidance and may have been responsible for highly consistent results in pilot test rankings.

\textsuperscript{19}The instructions provided to subjects are presented in Appendix B.
to three factors and typically each factor has three or four levels defined by the researcher. The axiomatic approach is applied to the data first. It entails testing the data to see if it complies with various properties of simple polynomial models. The objective is to posit a particular simple polynomial model by determining which properties do and do not hold. The four models that are considered are an additive model, a multiplicative model, a distributive model and a dual-distributive model. Explanation of the testing of the properties is provided in Appendix E. The flowchart (figure 8 of Appendix E) provides structure for understanding the relationship between the axioms and the simple polynomial models that are considered by CJM.

As seen in the flowchart (figure 8, Appendix E), simple independence for all three factors and double cancellation must be satisfied for any of the simple polynomials to be valid. Joint independence tests are used to distinguish between an additive and non-additive model. If all three pairs of factors (probability and effort, probability and consequence, and effort and consequence) satisfy joint independence, an additive model is concluded. If one or two pairs of factors satisfy joint independence, distributive cancellation and dual-distributive cancellation axioms are tested to determine if a distributive or dual-distributive model seems to fit the data better. If no pair of factors satisfies joint independence, no simple polynomial model is appropriate.

The numerical approach to conjoint measurement consists of deriving scales of measurement for the factors. The scales are determined such that the scaled factors, when combined according to an additive model, result in dependent variable values which reflect the subject's rank order of the combinations. The distance between the rank order achieved by the scaled values and the original rank ordered data is typically assessed by a
badness-of-fit measure. The scale values, called part worth functions, provide information about each factor's contribution to the auditor's preference for increased substantive testing.

5.2.4.1 Previous Applications of CJM in Accounting Research

CJM techniques have been used in two other areas of audit judgment. Moriarity and Barron (1976) used CJM to model the materiality judgments made by fifteen audit partners for eighteen situations, each with a different combination of factor levels. Schneider (1982, 1984) used CJM to model the external auditor's evaluation of the client's internal audit function.

Both of these previous studies concluded that the appropriate model to represent the audit judgment being made was an additive one. There are, however, two reasons why one could question an additive result. The first reason relates to the axiomatic conjoint measurement approach. Because data provided by subjects are rarely completely error free (although they appeared to be for eight of Moriarity and Barron's subjects), there will usually be some violations of the independence and double cancellation axioms that are required for an additive model. Most research using axiomatic CJM has used ad hoc acceptance levels for violations. For example, Schneider accepted as many as 20% violations of independence axioms and 30% violations of the double cancellation axiom before rejecting an additive model. Due to the lack of an error theory for dealing with violations of the axioms, additive models may have been erroneously assumed to be appropriate.

The second reason relates to the numerical approach. Nickerson and McClelland (1984) show that models generated from processes other than additively combining

\[20\text{One measure is STRESS resulting from Kruskal's monotonic transformation procedure (Kruskal 1965) and another is THETA which is related to Kendall's Tau coefficient (Johnson 1973).}\]
factors can be fit very well by an additive scaling procedure making an additive model
appear to be appropriate. They recommend that axiomatic CJM always be performed
first before assuming additive data and applying the numerical approach. Although
Schneider applied axiomatic CJM first, he supported the claim of an additive model, even
though there were axiomatic violations, by measuring the goodness (or badness) of fit to
an additive model using two approaches. It is possible that a non-additive model could
have generated the rank orderings even though numerical CJM resulted in a good-fitting
additive model (Emery and Barron 1979).

5.2.4.2 CMSCAL Applied to Audit Planning Judgments

The CMSCAL algorithm (Nygren 1985, 1986) eliminates some of the bias toward
concluding an additive model by analyzing the ordinal data in more detail than other
algorithms. The tests for simple independence and joint independence are further broken
down into tests for dominance and trade-off type violations. By identifying these two
types of violations, one can "differentiate between systematic and unsystematic errors in
fallible data" (Nygren 1985). If the errors are random, it would be appropriate to accept
a certain amount of error and conclude an additive model. Systematic errors point to a
non-additive model. So, although a limited number of unsystematic errors is tolerated
when CMSCAL is used, the distinction that is made between types of error eliminates
some of the ad hoc nature seen in previous applications of CJM analysis.

The use of CMSCAL therefore, allows analysis of the data provided by auditors
in this experiment at a greater level of detail than was available for previous applications
of CJM analysis to accounting research. If subjects attend to all three factors, it seems
reasonable to assume that simple independence should hold for each of the factors.
Simple independence would imply that high probability is preferred to medium
probability which is preferred to low probability, regardless of the associated combination
of effort and consequence. Similarly, low effort is expected to be preferred to medium effort which should be preferred to high effort, regardless of the associated combination of probability and consequence. Finally high consequence is expected to be preferred to medium consequence which should be preferred to low consequence, regardless of the associated combination of probability and effort. It seems that there should be no violations of simple independence in this auditing context. However, it is not clear what the axiom tests of joint independence will reveal. Satisfaction of the joint independence axiom for probability and consequence, for example, implies that if a particular combination of probability and consequence is preferred to another, it should always be preferred to the other, regardless of the associated level of the effort factor. The following section indicates what the implications of trade-off violations of joint independence axioms would be for audit planning decisions.

5.2.5 Examples of Trade-off Violations of Joint Independence Axioms

The main interest in using the CMSCAL algorithm is to be able to determine the nature of trade-offs made between the factors, if any. Both the distributive and dual-distributive models are consistent with trade-off violations of joint independence for one or two pairs of the factors at different levels of a third factor. It is conceivable in the audit planning context that there may be trade-offs between current and future expected cost and therefore, a non-additive model may be appropriate. The specific examples below illustrate what trade-off type violations would look like in this context. The interpretation of the trade-off type violations is more intuitive and provides more insight into the auditor's preferences than the specific non-additive model that is fit to the data. Appendix F illustrates each of the simple polynomial models using probability, effort and consequence as factors.
5.2.5.1 Probability and Effort

If probability and effort are not jointly independent of consequence, the following is an example of the preference relationships that could occur:

At high and medium levels of consequences:

\[ P_2 E_1 > P_1 E_3 \]

But, for low level of consequences:

\[ P_2 E_1 < P_1 E_3 \]

This pattern of preferences says that if the level of consequences is high or medium, an account having a medium level of probability and a high level of effort would be preferred to receive a high level of substantive testing over an account with a low level of probability and a low level of effort. When consequence is at a low level, however, the auditor would prefer to plan a high level of substantive testing for an account with a low level of probability and a low level of effort over an account with a medium level of probability and a high level of effort. This implies that the auditor is willing to trade-off probability for a reduction in effort if the consequences for not detecting an error are low.

5.2.5.2 Consequence and Effort

If consequence and effort are not jointly independent of probability, the following is an example of the preference relationships that could occur:

At high and medium levels of probability:

\[ C_2 E_1 > C_1 E_3 \]

But, for low level of probability:

\[ C_2 E_1 < C_1 E_3 \]

This pattern of preferences says that if the level of probability is high or medium, an account having a medium level of consequence and a high level of effort would be
preferred to receive a high level of substantive testing over an account with a low level of consequence and a low level of effort. When probability is at a low level, however, the auditor would prefer to plan a high level of substantive testing for an account with a low level of consequence and a low level of effort over an account with a medium level of consequence and a high level of effort. This implies that the auditor is willing to trade-off consequence for a reduction in effort if the probability of error is low.

5.2.5.3 Probability and Consequence

If probability and consequence are not jointly independent of effort, the following is an example of the preference relationships that could occur:

At low and medium levels of effort:

\[ P_1 C_2 > P_3 C_1 \]

But, for high level of effort:

\[ P_1 C_2 < P_3 C_1 \]

This pattern of preferences says that if the level of effort is low or medium, an account having a low level of probability and a medium level of consequence would be preferred to receive a high level of substantive testing over an account with a high level of probability and a low level of consequence. When effort is at a high level, however, the auditor would prefer to plan a high level of substantive testing for an account with a high level of probability and a low level of consequence over an account with a low level of probability and a medium level of consequence. This implies that the auditor is willing to trade-off probability and consequence, preferring less substantive testing for a higher consequence account in favor of emphasizing an account with a higher probability of error, if the effort required to increase the substantive testing is high.

There are many possible trade-off violations of the joint independence axiom. If subjects exhibit trade-off type violations, they can be interpreted in a manner similar to
the examples above to provide useful information about the preferences of auditors with regard to various combinations of characteristics represented by accounts. In addition to providing potential explanations to auditing researchers for prior experimental results, the information may be valuable to auditing firms.

5.3 The SAC Task and Questionnaire²¹

5.3.1 Development of the SAC Task

5.3.1.1 Background Information

Similar to previous auditing research, subjects were asked to read an auditing case and then make planning judgments. Case materials include a description of a client company and specific information intended to influence the subject's assessment of the probability of error in and the consequence level of the Accounts Receivable and Property, Plant and Equipment accounts.

A medium sized manufacturing company was chosen because most of the existing research, archival and experimental, is based on medium sized manufacturing firms. The availability of archival data and comparability to previous research were desired. Within the category of medium sized manufacturing companies, the criteria for a specific industry were that it should be one that is not an industry specialty of an auditing firm and not overly sensitive to economic cycles. The author therefore chose a company that makes underwater pumps and fractional horsepower motors. The company's annual report and 10K report were used to provide specific descriptions and to base financial statement relationships on those of an actual company.

²¹ Version A of the SAC task and the questionnaire are provided in Appendix C.
In contrast to other experimental audit planning research, two accounts were chosen as the focus of this experiment: accounts receivable (AR) and property, plant and equipment (PPE). The purpose of including two accounts was to investigate the auditor's planning decisions across accounts. A priori, it was believed that the two accounts would be perceived as having different consequence levels associated with them and that consequence level would have an impact on planning decisions.

The basic design of the case is similar to that of other researchers' materials (Libby 1985 and Bedard and Wright 1992). The type of information that was provided to subjects in the case was determined by considering other research instruments, audit manuals, planning documents and auditing texts.

Comparative financial statements (current year unaudited and previous two years audited), schedules for each of the accounts and comparative information over the three years regarding debt covenant provisions were provided. Three years of financial information were included so comparisons could be made between the current year amounts and those of the previous two years. Receivables schedules, one by age and one by size, were included to help in planning substantive tests for the AR and allowance account. PPE schedules were included to help in planning substantive tests for the PPE account. The previous year's audit hours were included to provide a benchmark for subjects to plan the current year's audit, consistent with previous research (Cohen and Kida 1989 and Bedard and Wright 1992).

Two versions, A and B, of the case were created. In version A, AR was designed to have a high probability of error, while PPE had a low probability of error. In version B, the description of the PPE account was intended to convey a high probability of error and the AR account had a low probability of error. In both versions, the accounts

---

22 Robert Libby and Arnold Wright each provided a copy of his research instrument.
receivable account was intended to be perceived as having a high consequence associated with not detecting an error therein. In both versions, PPE was intended to be perceived as having a lower consequence level than AR. Similar to previous research, the probability of error within each account was a between subjects factor, but both probability and consequence across the two accounts were within subjects factors.

5.3.1.2 Designing Levels of the Probability Factor

The probability of error for each account was operationalized by including some risk factors in the high probability version of the account that were not included in the low probability version of the same account. Quality of personnel was identified by Colbert (1988) as the most important factor for subjects in her experiment which asked auditors to assess the level of inherent risk in an inventory account. Personnel characteristics are also cited as risk factors in the auditing standards. Unexpected analytical review results have been shown to impact auditors' planning judgments (Cohen and Kida 1989). Therefore, the high probability of error descriptions for each account (AR in Version A and PPE in Version B) included a significant difference from the dollar amount that might be expected in each account and indicated personnel problems associated with the accounting function related to each account. The characteristics used to influence perceived error probability are further supported by the archival error literature discussed in Chapter II.

5.3.1.3 Designing Levels of the Consequence Factor

The difference in consequence level between the two accounts was operationalized, in part, through debt covenant provisions. In both Versions A and B, one of the debt covenant provisions was the current ratio. Since the net accounts receivable is included in the numerator of the current ratio, a misstatement of accounts
receivable could inflate (or deflate) the current ratio. Both versions showed the same pattern for the current ratio over the three years presented. In 1990, the current ratio was 1.88 and in 1991, it increased to 1.89. The required current ratio was stated at greater than or equal to 1.60. The 1992 unaudited current ratio was given to be 1.65. The ratio was shown as declining significantly in the current year, approaching the minimum required current ratio.\textsuperscript{23} In conversations with the consulting auditors, most indicated that proximity to debt covenant provisions would be a factor in evaluating the level of consequences for an account. The PPE account was not involved in any of the debt covenant requirements.

In addition to the debt covenant provision, the nature of the AR account is assumed to imply a higher level of consequence. Archival research (see section 2.3) shows that fraud is more prevalent in AR. Palmrose's (1987) litigation data indicate a higher incidence of auditor payments when management fraud is involved in an alleged audit failure. The debt covenant provision and the greater susceptibility of AR to fraud were intended to create a perception of more severe consequences associated with not detecting a material error in AR, compared to not detecting a material error in PPE.

5.3.2 Dependent Variables

Some of the planning judgments that the subjects were asked to make differed from what other researchers have asked in similar experiments. One objective of this research is to explore the factors that influence auditors' planning judgments for different accounting situations.

\textsuperscript{23}In Version A, where there is a high probability of error in the AR account, it is likely that the client could actually be in technical default on the debt due to a violation of the debt covenant. The probability of error and the consequence level may interact in Version A to result in different judgments being made than what would be suggested by judgments made at lower consequence and/or probability levels.
accounts within an audit. Other research has included only one account. By including two accounts, this study may find that the planning judgment process differs between accounts.

Other research has asked only for the types of tests and number of audit hours, as an indication of the level of testing that would be planned. Asking only for audit hours as an indication of audit intensity may be misleading. Using a within subjects design, Joyce (1976) showed that different auditors can have very different ranges of audit hours for the same set of situations. Each subject in Joyce's experiment provided an audit hour estimate for each of 20 or 36 cases. The average number of hours a subject estimated over all situations ranged from 18 to 120 hours. Joyce (1976) was able to determine how each case was perceived by an auditor by comparing the number of hours planned for one case to that auditor's average number of audit hours. But most audit planning research has used a between subjects design, where each subject is asked to estimate the number of audit hours for one account in one situation. It is not clear how a particular situation is perceived by the auditor, because each may have a different scale of the possible number of hours he or she would plan. Two auditors could agree that a situation merits a particular level of detection risk but, because of differences in auditing skills or styles, could estimate very different numbers of audit hours. The questions asked in this research, in addition to an estimate of audit hours, attempt to sort out some of the diversity seen in previous research when only audit hours were requested.

After reading the case material, the first question asked of the subject was to indicate his or her relative level of concern for each of the accounts. This measure of "concern" was intended to elicit the auditor's relative focus on the accounts within the audit. Because of the nature of the two accounts, it was not reasonable to expect that the relative audit hours would necessarily represent the auditor's relative concern for the accounts. The AR account comprises many more transactions than the PPE account,
requiring more audit resources to perform substantive testing even if only a low level of
testing is intended for AR. The term "concern" was used because it is not a technical
auditing term. The ratings were intended to result in an overall assessment of each
account by allowing each individual to incorporate what he or she worries about when
considering how to plan testing of the account. Using a technical term, such as "risk"
may have conjured up different specific images for each auditor.

After considering the two accounts together, the subject was then asked to
consider each of the accounts separately. First he or she was asked to specify, on a scale
of -5 to +5, the level of substantive testing he or she would plan for the AR account for
this audit. Level "-5" ("+5") on the scale was described as "Minimum (Maximum) level
of testing for comparably sized client." The "0" point was described as "Average level of
testing for comparably sized client." The purpose of this scale was to determine how the
AR account described in the case compared with other AR situations the auditor may
have seen. Rather than trying to infer the auditors' relative levels of emphasis from audit
hours, the scale was intended to provide an indication of how planned audit hours for the
described account differ from what the auditor would plan for the same account under
different circumstances.

Because the amount of reliance the audit firm decides to place on the internal
control system can have a significant impact on the level of substantive testing that is
planned (O'Keefe, Simunic and Stein 1992 and Cohen and Kida 1989), it was important
that the account be compared to other AR accounts in a similar reliance situation. Since
the case stated that the audit firm had decided not to rely on the internal control system
due to the high cost of testing controls, the subjects were instructed to make comparisons
on that basis.
After providing a rating for the level of testing to be done, the subject was asked to indicate which substantive audit procedures he or she would plan for the AR account and indicate the total number of hours required to audit the account. The subject was provided with a list of procedures that were compiled by the researcher from several sources (auditing texts, the audit manual of one of the Big Six firms and an audit program from a large regional auditing firm). The subject could check off procedures from the list and add any other procedures he or she felt were warranted. These requirements are comparable to those in Joyce (1976), Cohen and Kida (1989) and Bedard and Wright (1992).

The subject was then asked to provide a rating and an estimate of the number of audit hours for planned procedures for the PPE account in a similar manner to what was done for AR. Subjects provided an indication of the planned level of substantive testing and the approximate number of hours he or she would plan for the audit of each account.

5.3.3 Post-Experimental Questionnaire

The purpose of the questionnaire was to elicit biographical information and assessments from the subjects. In order to make sure that the probability and consequence levels were perceived as intended, the questionnaire asked the subject to provide an assessment of the probability and consequence levels for both accounts. The level of perceived consequences consisted of both a personal aspect and a firm aspect. By asking about both aspects we may learn whether it is the consequences for the auditor personally or the consequences for the firm, that influences planning judgments. These assessed levels of probability and consequences were used as independent variables in regression models attempting to explain variation in levels of concern for the accounts, planned levels of substantive testing, and planned audit hours.
Although the order of the two tasks was alternated randomly for the subjects, the questionnaire was always included with the SAC task, whether it was the first or second task. Because the probability and consequence assessments were in the questionnaire, it was important to ask those questions right after the SAC task was completed.

5.3.4 Pilot Testing of the SAC Task

The SAC task was sent to eleven of the CJM pilot subjects. Ten responses were received and of those, nine were complete. In general, the pilot test results were consistent with expectations.
CHAPTER VI
RESULTS

The first section of this chapter describes the auditors who participated in the study. Section 6.2 presents results of manipulation checks testing the intended design of the SAC task. Section 6.3 reports the results of full regression models for planning judgments. Sections 6.4 through 6.8 report the results of tests of the research hypotheses. Section 6.9 discusses the results of the CJM task.

6.1 Descriptive Data

Table 4 on the next page provides information about the sixty-five auditors who participated in this study.24 Five of the Big Six auditing firms and one large regional firm are represented by the subjects. More than half (33) of the sixty-five were seniors and the years of audit experience for the entire group ranged from eight months to eight and one-half years. Planning experience for each account was measured by the number of audits for which the subjects had planned substantive tests for either AR or PPE. Four auditors had not planned substantive tests for any AR account and four had not planned substantive tests for any PPE account.25 Three of these four auditors had not planned substantive tests for either account.

---

24 Because one of the subjects did not complete the SAC task and the questionnaire only 64 subjects are represented in panels B through D.

25 The results of analyses that excluded these subjects were not qualitatively different from the results of analyses which included these subjects. Therefore, reported results include all subjects.
Table 4. Participating Auditors, Descriptive Data

Panel A: Firm Affiliation

<table>
<thead>
<tr>
<th>Firm</th>
<th>Firm Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>

Total Auditors: 65

Panel B: Years of Audit Experience

<table>
<thead>
<tr>
<th>Experience</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>.67</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>8.50</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>3.38</td>
<td>1.52</td>
</tr>
</tbody>
</table>

Panel C: Planning Experience

<table>
<thead>
<tr>
<th>Number of Audits Planned</th>
<th>Total Auditors</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1-3</td>
<td></td>
</tr>
<tr>
<td>4-7</td>
<td></td>
</tr>
<tr>
<td>8-12</td>
<td></td>
</tr>
<tr>
<td>13-18</td>
<td></td>
</tr>
<tr>
<td>&gt;18</td>
<td></td>
</tr>
</tbody>
</table>

AR

<table>
<thead>
<tr>
<th>Number of Audits Planned</th>
<th>Total Auditors</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>64</td>
</tr>
<tr>
<td>19</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

PPE

<table>
<thead>
<tr>
<th>Number of Audits Planned</th>
<th>Total Auditors</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>64</td>
</tr>
<tr>
<td>17</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

Panel D: Rank

<table>
<thead>
<tr>
<th>Rank</th>
<th>Total Auditors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff</td>
<td>14</td>
</tr>
<tr>
<td>Senior</td>
<td>33</td>
</tr>
<tr>
<td>Supervisor</td>
<td>14</td>
</tr>
<tr>
<td>Manager</td>
<td>3</td>
</tr>
</tbody>
</table>

Total Auditors: 64
Subjects were randomly assigned to either Version A or Version B of the SAC task. Thirty-four subjects completed Version A and 30 subjects completed Version B. Table 5 illustrates that there were no significant differences in the years of auditing experience ($p = .84$), planning experience for AR ($p = .62$) or planning experience for PPE ($p = .83$) between groups.

Table 5. Descriptive Statistics - Version A versus Version B Subjects

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>T-statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of Experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Version A</td>
<td>3.42</td>
<td>1.52</td>
<td>-.21</td>
<td>.84</td>
</tr>
<tr>
<td>Version B</td>
<td>3.34</td>
<td>1.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning Experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Version A</td>
<td>8.35</td>
<td>6.91</td>
<td>-.49</td>
<td>.62</td>
</tr>
<tr>
<td>Version B</td>
<td>7.53</td>
<td>6.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning Experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Version A</td>
<td>8.46</td>
<td>6.84</td>
<td>-.22</td>
<td>.83</td>
</tr>
<tr>
<td>Version B</td>
<td>8.10</td>
<td>6.19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

26 Within each firm, the number of participants who completed each version was about equal, except for Firm 6, where four subjects completed Version A and one subject completed Version B. Although the materials distributed to this firm consisted of equal numbers of Version A and Version B, the response rate for this firm was very low, resulting in an uneven distribution of versions in the responses.
6.2 Results of Manipulation Checks

As explained in section 1.3.2, this research differs from previous experimental audit planning research in that PPE is included as well as AR. It was expected that AR would be perceived as having a higher level of consequence associated with not finding an existing error compared to not finding an existing error in PPE. It was also expected that the perceived level of probability of error could be manipulated between accounts. In Version A of the SAC task, AR was intended to have a higher probability of error than PPE. In Version B, PPE was intended to have a higher probability of error than AR.

The intended levels of probability and consequence for the two accounts in both versions were as illustrated in figure 4.

<table>
<thead>
<tr>
<th>Accounts Receivable</th>
<th>Property, Plant and Equipment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Version A</strong></td>
<td></td>
</tr>
<tr>
<td>Consequence = H or M</td>
<td>Consequence = L</td>
</tr>
<tr>
<td>Probability of error = H</td>
<td>Probability of error = L</td>
</tr>
<tr>
<td><strong>Version B</strong></td>
<td></td>
</tr>
<tr>
<td>Consequence = H or M</td>
<td>Consequence = L</td>
</tr>
<tr>
<td>Probability of error = L</td>
<td>Probability of error = H</td>
</tr>
</tbody>
</table>

Where L - Low, M - Medium and H - High

**Figure 4. Design of SAC Task**

6.2.1 Manipulation Checks Related to Consequence levels

It was expected that the assessed consequence level associated with not detecting a material error in AR would be significantly higher than the assessed consequence level associated with not detecting a material error in PPE. Over both versions of the SAC task, the average consequence level for AR was 6.34 (on a scale of 1 to 10) and the

27 The "average consequence level" in this context consists of a simple average of a subject's personal and firm consequence ratings.
average consequence level for PPE was 5.73. A paired t-test indicated that the mean difference between subjects' consequence assessments for the two accounts was significantly different from zero (panel A of table 6). Panel B of table 6 shows the two consequence assessments that were made by subjects and the results of paired t-tests. The assessed consequence levels (both personal and firm) related to AR were significantly higher than the assessed consequence levels related to PPE in Version A of the SAC task. In Version B where there were more risk factors related to PPE, the assessed personal consequence level for AR was only slightly higher than the personal consequence level for PPE. The assessed firm consequence level for AR was significantly higher than the assessed firm consequence level for PPE.

The results of the paired t-tests indicate that on average, the subjects assessed the consequences associated with not finding an error in the AR account as higher than the consequences associated with not finding an error in the PPE account. The difference between personal consequence levels was not statistically significant for those who completed version B of the SAC task.
Table 6. Results of Manipulation Checks for Consequence

Panel A: Paired T-test for Average Consequence Assessments - Overall

<table>
<thead>
<tr>
<th></th>
<th>AR Mean</th>
<th>PPE Mean</th>
<th>Paired t-statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.34</td>
<td>5.73</td>
<td>4.50</td>
<td>.00</td>
</tr>
</tbody>
</table>

Panel B: Paired T-tests for Consequence Assessments by Version of SAC Task

<table>
<thead>
<tr>
<th></th>
<th>Consequence (Personal)</th>
<th>Consequence (Firm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VERSION A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR Mean</td>
<td>5.94</td>
<td>7.18</td>
</tr>
<tr>
<td>PPE Mean</td>
<td>5.23</td>
<td>6.15</td>
</tr>
<tr>
<td>Paired t-statistic</td>
<td>3.85</td>
<td>5.75</td>
</tr>
<tr>
<td>P-value</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>VERSION B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR Mean</td>
<td>4.98</td>
<td>7.20</td>
</tr>
<tr>
<td>PPE Mean</td>
<td>4.92</td>
<td>6.63</td>
</tr>
<tr>
<td>Paired t-statistic</td>
<td>.30</td>
<td>2.01</td>
</tr>
<tr>
<td>P-value</td>
<td>.76</td>
<td>.05</td>
</tr>
</tbody>
</table>
6.2.2 Manipulation Checks Related to Probability Levels

It was expected that the assessed probability of error for AR would be significantly higher than the assessed probability of error for PPE in Version A. Panel A of table 7 shows the results of paired t-tests comparing the assessed probability levels within versions of the SAC task. The assessed probability level for AR was 6.24 and the assessed probability level for PPE was 3.02 in Version A. The difference between the accounts' assessed probability levels in Version A is statistically significant and in the expected direction.

It was expected that the assessed probability of error for PPE would be significantly higher than the assessed probability of error for AR in Version B of the SAC task. Panel A of table 7 shows the results of paired t-tests comparing the assessed probability levels. The assessed probability level for PPE was 4.79 and the assessed probability level for AR was 3.58 in Version B. The difference between the accounts' assessed probability levels in Version B is statistically significant and in the expected direction.

It was expected that the assessed probability of error for AR would be significantly higher in Version A of the SAC task than in Version B. Panel B of table 7 shows the result of an independent t-test comparing the assessed probability levels of AR in Version A and Version B. The assessed probability level for AR in Version A was 6.24 and 3.58 in Version B. The difference between probability assessments for AR across versions is statistically significant and in the expected direction.
Table 7. Results of Manipulation Checks for Probability

Panel A: Paired T-tests of Probability Assessments

<table>
<thead>
<tr>
<th></th>
<th>AR Mean</th>
<th>PPE Mean</th>
<th>Paired t-statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VERSION A</td>
<td>6.24</td>
<td>3.02</td>
<td>10.50</td>
<td>.00</td>
</tr>
<tr>
<td>VERSION B</td>
<td>3.58</td>
<td>4.79</td>
<td>-2.89</td>
<td>.01</td>
</tr>
</tbody>
</table>

Panel B: Independent T-tests of Probability Assessments

<table>
<thead>
<tr>
<th>Probability of Error - AR</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>T-statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version A</td>
<td>6.24</td>
<td>1.47</td>
<td>-7.42</td>
<td>.00</td>
</tr>
<tr>
<td>Version B</td>
<td>3.58</td>
<td>1.41</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Probability of Error - PPE</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>T-statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version A</td>
<td>3.02</td>
<td>1.32</td>
<td>3.95</td>
<td>.00</td>
</tr>
<tr>
<td>Version B</td>
<td>4.79</td>
<td>2.12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
It was expected that the assessed probability of error for PPE would be significantly higher in Version B of the SAC task than in Version A. Panel B of table 7 shows the result of an independent t-test comparing the assessed probability levels of PPE in Version A and Version B. The assessed probability level for PPE in Version B was 4.79 and 3.02 in Version A. The difference between probability assessments for PPE across versions is statistically significant and in the expected direction.

6.2.3 Summary of Results Related to Manipulation Checks

The results suggest that subjects reacted as expected to the risk factors included in each version when assessing probabilities of error in each account. For the most part, the consequence assessments were consistent with expectations. It was expected that AR would always be perceived as having a higher consequence associated with it compared to PPE. The pattern of assessed consequences for the firm were more consistent with expectations than the pattern of assessed personal consequences. The mean consequence level for the firm was approximately 7.2 for AR in both versions of the SAC task. The mean consequence level for the firm associated with the PPE account increased from Version A to Version B. It appears that the increase in probability of error for PPE was associated with an increase in perceived consequences for the firm related to PPE. The personal consequence assessments, however, decreased for both AR and PPE in going from Version A to Version B.
6.3 Regression Model for Planning Judgments

A comprehensive regression model was used with six different dependent variables to test the effects of probability, consequence, type, order and appropriate interactions. The set of independent variables was expected to provide significant explanatory power for 1) the level of concern for the Accounts Receivable (AR) account, 2) the level of concern for the Property, Plant and Equipment (PPE) account, 3) the level of substantive testing for the AR account, 4) the level of substantive testing for the PPE account, 5) total planned audit hours for AR and 6) total planned audit hours for PPE. The full model and definitions of the variables are shown in figure 5. Table 8 reports the results of the regression analysis for all six models. Sections 6.4 through 6.8 discuss the results with respect to each hypothesis (H1 through H6) and present additional tables where necessary.
\[ LC_{ij} (LST_{ij}, AH_{ij}) = \beta_0 + \beta_1 (\text{Prob}_{ij}) + \beta_2 (\text{Cons}_{ij}) + \beta_3 (\text{Prob}_{ij} \times \text{Cons}_{ij}) + \beta_4 (\text{TypeP}_i) + \beta_5 (\text{TypeC}_i) + \beta_6 (\text{TypeP}_i \times \text{Prob}_{ij}) + \beta_7 (\text{TypeC}_i \times \text{Cons}_{ij}) + \beta_8 (\text{Order}_i) + \beta_9 (\text{Order}_i \times \text{Prob}_{ij}) + \beta_{10} (\text{Order}_i \times \text{Cons}_{ij}) + \epsilon_{ij} \]

Where:

\( LC_{ij} \) = subject i’s rating of the level of concern he or she has for account j as it is described in the case materials for the hypothesized client.

Range of the variable: 1 (very low concern account) to 10 (very high concern account)

\( LST_{ij} \) = subject i’s rating of the level of substantive testing he or she would plan for account j compared to what he or she would plan for the same account in other circumstances.

Range of the variable: -5 ("minimum level of testing for comparably sized client") to +5 ("maximum level of testing for comparably sized client"). The range was converted to a 1 to 11 range for the actual data analysis to make the values of the variable comparable to the LC variable.

\( AH_{ij} \) = the total estimated audit hours planned by subject i for account j

\( \text{Prob}_{ij} \) = subject i’s rating of the probability (likelihood) of a material error in account j

Range of the variable: 0 (no chance) to 10 (certain)

\( \text{Cons}_{ij} \) = subject i’s average of the personal and firm consequence ratings associated with not detecting an existing error in account j during the audit

Range of variable: 1 (not very severe at all) to 10 (very severe)

\( \text{Prob}_{ij} \times \text{Cons}_{ij} \) = probability x consequence interaction

Range of variable: 0 to 100

\( \text{TypeP}_i \) = indicator variable, = 1 if subject i indicates by his or her rankings in the CJM task that probability is clearly the most important factor, = 0 otherwise.

\( \text{TypeC}_i \) = indicator variable, = 1 if subject i indicates by his or her rankings in the CJM task that consequence is clearly the most important factor, = 0 otherwise.

\( \text{TypeP}_i \times \text{Prob}_{ij} \) = TypeP x probability interaction.

Range of variable: 0 to 10

\( \text{TypeC}_i \times \text{Cons}_{ij} \) = TypeC x Consequence interaction

Range of variable: 0 to 10

Figure 5. Full Regression Model and Definitions of Variables
Figure 5 (continued)

\textbf{Order}_i \equiv \text{An indicator variable that indicates the order in which subject } i\text{ completed the two tasks, } = 1 \text{ if subject } i\text{ did the CJM task first, } = 0 \text{ if subject } i\text{ did the Specific Audit Scenario task first.}

\textbf{Order}_i \times \text{Prob}_i \equiv \text{Order } \times \text{Probability interaction}
\begin{itemize}
  \item \textbf{Range of variable:} 0 \text{ to } 10
\end{itemize}

\textbf{Order}_i \times \text{Cons}_i \equiv \text{Order } \times \text{Consequence interaction}
\begin{itemize}
  \item \textbf{Range of variable:} 0 \text{ to } 10
\end{itemize}
Table 8. Full Regression Models - Results

<table>
<thead>
<tr>
<th></th>
<th>LC_AR Coeff.</th>
<th>LST_AR Coeff.</th>
<th>AH_AR Coeff.</th>
<th>LC_PPE Coeff.</th>
<th>LST_PPE Coeff.</th>
<th>AH_PPE Coeff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>B₀</td>
<td>-.68 (0.81)</td>
<td>3.44 (0.22)</td>
<td>18.68 (0.56)</td>
<td>4.32** (0.03)</td>
<td>6.84*** (0.00)</td>
<td>19.85* (0.07)</td>
</tr>
<tr>
<td>B₁ (Prob)</td>
<td>1.39*** (.00)₁</td>
<td>.62* (.10)₁</td>
<td>4.33 (.22)₁</td>
<td>.60** (.05)₁</td>
<td>.17 (.28)₁</td>
<td>-1.58 (0.78)₁</td>
</tr>
<tr>
<td>B₂ (Cons)</td>
<td>.65* (.06)₁</td>
<td>.13 (.37)₁</td>
<td>1.59 (.37)₁</td>
<td>-.22 (.77)₁</td>
<td>-.15 (.73)₁</td>
<td>-1.90 (0.87)₁</td>
</tr>
<tr>
<td>B₃ (Prob x Cons)</td>
<td>-.09 (.91)₁</td>
<td>.02 (.42)₁</td>
<td>-.17 (.59)₁</td>
<td>-.02 (.64)₁</td>
<td>.01 (.40)₁</td>
<td>.40 (0.13)₁</td>
</tr>
<tr>
<td>B₄ (TypeP)</td>
<td>1.45 (.20)</td>
<td>-.91 (.41)₁</td>
<td>17.35 (.17)</td>
<td>-.36 (.66)₁</td>
<td>.03 (.96)</td>
<td>5.44 (0.25)</td>
</tr>
<tr>
<td>B₅ (TypeC)</td>
<td>-.03 (.99)</td>
<td>1.28 (.45)₁</td>
<td>6.20 (.75)</td>
<td>-2.15* (.09)</td>
<td>-.69 (.50)</td>
<td>-3.61 (0.62)</td>
</tr>
<tr>
<td>B₆ (TypeP x Prob)</td>
<td>-.25 (.89)₁</td>
<td>.12 (.27)₁</td>
<td>-.31 (.92)₁</td>
<td>-.12 (.77)₁</td>
<td>-.13 (.82)₁</td>
<td>-1.58 (0.94)₁</td>
</tr>
<tr>
<td>B₇ (TypeC x Cons)</td>
<td>.00 (.50)₁</td>
<td>-.23 (.81)₁</td>
<td>1.59 (.70)₁</td>
<td>.32* (.06)₁</td>
<td>.14 (.20)₁</td>
<td>1.27 (0.13)₁</td>
</tr>
<tr>
<td>B₈ (Order)</td>
<td>1.12 (.51)</td>
<td>.23 (.89)₁</td>
<td>-2.24 (.91)</td>
<td>-2.09* (.10)</td>
<td>-1.58 (.13)</td>
<td>-10.60 (0.14)</td>
</tr>
<tr>
<td>B₉ (Order x Prob)</td>
<td>-.20 (.34)</td>
<td>-.14 (.50)₁</td>
<td>1.13 (.63)</td>
<td>.26 (.13)</td>
<td>.31** (.03)</td>
<td>2.40** (.02)</td>
</tr>
<tr>
<td>B₁₀ (Order x Cons)</td>
<td>.03 (.91)</td>
<td>.08 (.76)₁</td>
<td>1.51 (.62)</td>
<td>.34* (.09)</td>
<td>.02 (.93)</td>
<td>.78 (0.50)</td>
</tr>
</tbody>
</table>

*p significant at α = .10, **p significant at α = .05, ***p significant at α = .01

28 Reported p-values having a subscript 1 are for one-sided tests, other p-values are for two-sided tests.
Table 8 (continued)

<table>
<thead>
<tr>
<th></th>
<th>LCA_R</th>
<th>LSTAR</th>
<th>AHR</th>
<th>LPPE</th>
<th>LSTPPE</th>
<th>AH_PPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted $R^2$</td>
<td>.42</td>
<td>.50</td>
<td>0</td>
<td>.52</td>
<td>.37</td>
<td>.32</td>
</tr>
</tbody>
</table>
6.4 Results Related to Tests of Hypothesis One

Hypothesis one (H1) predicts that an increase in probability of error in an account will be associated with an increase in each of the substantive testing dependent variables. With respect to the full regression model defined in table 8, H1 predicts that $\beta_1$ will be significantly greater than zero in each of the models. H1 is supported in the full regression model for AR when the level of concern ($\beta_1$ equals 1.39 and $p = .00$) or level of substantive testing ($\beta_1 = .62$ and $p = .10$) is used as the dependent variable, but not when the number of audit hours is used ($p = .22$). H1 is supported for PPE when level of concern ($\beta_1 = .60$ and $p = .05$) is used as the dependent variable, but not when the relative level of substantive testing ($p = .28$) or number of audit hours is used ($p = .78$).

When the level of substantive testing or number of audit hours for PPE is used as the dependent variable, the probability by order interaction is significant. This result implies that the probability assessment had more impact on substantive test levels for subjects who completed the CJM task first compared to those who completed the SAC task first. The following two subsections report results of reduced models which further test for the significance of the probability variable to substantive testing judgments.
6.4.1 Reduced Probability Models for AR

In addition to testing the effect of assessed probability of error in the full regression model, reduced models were tested for the AR account. These reduced models included probability as the only independent variable. Table 9 below shows the results of the reduced models for AR. The results indicate that assessed probability of error is able to explain 44 percent of the variation in level of concern for AR and about 56 percent of the variation in relative level of substantive testing for AR. Assessed probability of error explains only about 4 percent of variation in the number of audit hours that would be planned for AR. These results are comparable to the adjusted $R^2$s in the corresponding full models. Probability alone explains more variation in the dependent variables than all the dependent variables together. This indicates that probability is most likely the most significant variable of all those considered for AR.

Table 9. Reduced Probability Models for AR - Results

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Coefficients (p-value)$^{29}$</th>
<th>$R^2$ for the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant Probability</td>
<td></td>
</tr>
<tr>
<td>LCA_AR</td>
<td>4.52*** (.00)</td>
<td>.60*** (.00)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.44</td>
</tr>
<tr>
<td>LST_AR</td>
<td>4.14*** (.00)</td>
<td>.71*** (.00)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.56</td>
</tr>
<tr>
<td>AHA_AR</td>
<td>37.56*** (.00)</td>
<td>1.45* (.06)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.04</td>
</tr>
</tbody>
</table>

*$^p$ significant at $\alpha = .10$, $^{***}p$ significant at $\alpha = .01$

$^{29}$P-values for the probability variable are one-sided. Two-sided p-values are reported for the constant.
6.4.2 Reduced Probability Models for PPE

In addition to testing the effect of assessed probability of error in the full regression model, reduced models were tested for the PPE account. The reduced model for level of concern included probability as the only independent variable. Panel A of table 10 reports results of the reduced model for \( LC_{PPE} \). Probability of error was able to explain about 41 percent of the variation in level of concern for PPE.

Since the probability by order interaction was a significant variable in the other two full models for PPE, the reduced models for \( LST_{PPE} \) and \( AH_{PPE} \) which include probability also include order and the probability by order interaction. Panel B of table 10 shows the results of these reduced models.

The reduced models for the relative level of substantive testing for PPE indicate that probability, order and the probability by order interaction are all significant explanatory variables. These three independent variables together explain about 35 percent of the variation in relative testing levels for PPE. The results imply that subjects who completed the CJM task first start out with lower levels of substantive testing for PPE, but increase substantive testing more for increases in probability than subjects who completed the SAC task first.

The reduced model for PPE audit hours indicates that probability is not a significant explanatory variable but order and the probability by order interaction do provide significant explanation for the number of audit hours. These three independent variables together explain about 26 percent of the variation in audit hours for PPE. The impact of the two significant coefficients is similar to those in the relative level of substantive testing reduced model. The coefficients suggest that those who completed the CJM task would plan for a much smaller number of audit hours for PPE in the absence of probability of error, but would increase the number of audit hours at a faster rate than others as the probability of error increases.
Table 10. Reduced Probability Models for PPE - Results

Panel A: Reduced Model Including Probability Only

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Coefficients (p-value)</th>
<th>R² for the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant Probability</td>
<td></td>
</tr>
<tr>
<td>LC&lt;sub&gt;PPE&lt;/sub&gt;</td>
<td>2.53*** (.00)</td>
<td>.57*** (.00)</td>
</tr>
</tbody>
</table>

Panel B: Reduced Models Including Probability and Order Variables

\[ LST, AH = \beta_0 + \beta_1 (\text{Prob}_i) + \beta_2 (\text{Order}_i) + \beta_3 (\text{Order}_i \times \text{Prob}_i) + \varepsilon_i \]

<table>
<thead>
<tr>
<th>Coefficient (p-value)</th>
<th>Coefficient (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>\beta_0</td>
<td>6.00*** (.00)</td>
</tr>
<tr>
<td>\beta_1</td>
<td>.19** (.03)</td>
</tr>
<tr>
<td>\beta_2</td>
<td>-1.41** (.01)</td>
</tr>
<tr>
<td>\beta_3</td>
<td>.30** (.02)</td>
</tr>
</tbody>
</table>

Adjusted R² for the Model

* p significant at \( \alpha = .10 \), ** p significant at \( \alpha = .05 \), *** p significant at \( \alpha = .01 \)

Reported p-values having a subscript 1 are for one-sided tests, other p-values are for two-sided tests.
6.4.3 Summary of the Effect of Probability

Overall, the assessed probability level seemed to impact subject's audit planning judgments for AR as expected except when the number of audit hours was estimated. The inability of assessed probability of error to provide significant explanation for the number of audit hours planned is consistent with the results of related auditing research (specifically Bedard and Wright 1992).

For PPE, probability of error was able to explain significant variation in the level of concern for the account. Probability was also a significant explanatory variable in a reduced model for $L_{ST_{PPE}}$. It seems, however, that completing the CJM task first increased the sensitivity of auditors to the impact of the probability of error on substantive testing levels for PPE. Subjects who completed the CJM task first tended to plan lower relative levels of substantive testing for the PPE account if probability of error was low, but increased their relative testing levels more than other subjects in response to increased levels of probability of error. Consistent with the AR results, probability of error was not a significant variable in the models which explained variation in the number of audit hours planned for PPE.

6.5 Results of Tests of Hypothesis Two

Hypothesis two (H2) predicts that an increase in the consequence of not detecting a material error in an account will be associated with an increase in each of the substantive testing dependent variables. With respect to the full regression model defined in table 8, H2 predicts that $\beta_2$ will be significantly greater than zero in each of the models. H2 is supported in the full regression model for AR when the level of concern ($\beta_2$ equals .65 and $p = .06$) is used as the dependent variable, but not when the relative
level of substantive testing ($p = .37$) or the number of audit hours is used ($p = .37$). H2 is not supported for PPE when level of concern ($p = .77$), relative level of substantive testing ($p = .73$) or number of audit hours ($p = .87$) is used as the dependent variable.

When the level of concern for PPE is used as the dependent variable, the consequence by order interaction and the type by consequence interaction are significant. This result implies that the consequence assessment had more impact on substantive test levels for subjects who completed the CJM task first compared to those who completed the SAC task first. It also implies that those who indicated that consequence is the most important factor increased their concern for PPE more than others as the consequence level increased. The following two subsections report results of reduced models which further test for the significance of the consequence variable to substantive testing judgments.

6.5.1 Reduced Consequence Models for AR

In addition to testing the effect of average assessed consequence level in the full regression model, reduced models were tested for the AR account. Each set of reduced models included one of the consequence variables (personal, firm and average) as the only independent variable. Table 11 shows the results of the reduced models for AR. The results in panel A indicate that assessed level of personal consequences is able to explain 12 percent of the variation in level of concern for AR and about 13 percent of the variation in relative level of substantive testing for AR. The personal consequence coefficient is significantly greater than zero in the LC model ($p = .00$) and in the LST model ($p = .00$). Assessed personal consequence level is not a significant explanatory variable for the number of audit hours that would be planned for AR ($p = .19$).

The results in panel B indicate that assessed level of firm consequences is able to explain about 11 percent of the variation in the level of concern for AR and about 5
percent of the variation in the relative level of substantive testing for AR. The firm consequence coefficient is significantly greater than zero in the LC model \((p = .01)\) and in the LST model \((p = .03)\). Assessed firm consequence level is not a significant explanatory variable for the number of audit hours that would be planned for AR \((p = .17)\).

The results in panel C indicate that the average assessed consequence level is able to explain about 17 percent of the variation in the level of concern for AR and about 13 percent of the variation in the relative level of substantive testing for AR. The average consequence coefficient is significantly greater than zero in the LC model \((p = .00)\) and in the LST model \((p = .00)\). The average assessed consequence level is not a significant explanatory variable for the number of audit hours that would be planned for AR \((p = .13)\).
Table 11. Reduced Consequence Models for AR - Results

Panel A: Reduced Models Including the Personal Consequence Assessment

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Coefficients (p-value)</th>
<th>R² for the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant</td>
<td>Personal Conseq.</td>
</tr>
<tr>
<td>LC_ISR</td>
<td>5.87*** (.00)</td>
<td>.30*** (.00)</td>
</tr>
<tr>
<td>LST_ISR</td>
<td>5.92*** (.00)</td>
<td>.33*** (.00)</td>
</tr>
<tr>
<td>AH_ISR</td>
<td>40.38*** (.00)</td>
<td>.80 (.19)</td>
</tr>
</tbody>
</table>

Panel B: Reduced Models Including the Firm Consequence Assessment

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Coefficients (p-value)</th>
<th>R² for the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant</td>
<td>Firm Consequence</td>
</tr>
<tr>
<td>LC_ISR</td>
<td>5.01*** (.00)</td>
<td>.35*** (.01)</td>
</tr>
<tr>
<td>LST_ISR</td>
<td>5.83*** (.00)</td>
<td>.26** (.03)</td>
</tr>
<tr>
<td>AH_ISR</td>
<td>37.13*** (.00)</td>
<td>1.06 (.17)</td>
</tr>
</tbody>
</table>

31 Reported p-values for the consequence variables are for one-sided tests. Two-sided p-values are reported for the constant.
Table 11 (continued)

Panel C: Reduced Models Including the Average Consequence Assessment

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Coefficients (p-value)</th>
<th>R² for the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant</td>
<td>Average Conseq.</td>
</tr>
<tr>
<td>LC_{AR}</td>
<td>4.47*** (.00)</td>
<td>.48*** (.00)</td>
</tr>
<tr>
<td>LST_{AR}</td>
<td>4.84*** (.00)</td>
<td>.45*** (.00)</td>
</tr>
<tr>
<td>AH_{AR}</td>
<td>36.14*** (.00)</td>
<td>1.36 (.13)</td>
</tr>
</tbody>
</table>

**p significant at \( \alpha = .05 \), ***p significant at \( \alpha = .01 \)
6.5.2 Reduced Consequence Models for PPE

In addition to testing the effect of average consequence level in the full regression model, reduced models were tested for the PPE account. Since the consequence by order interaction and the consequence by type interaction were significant in the LCPPE full model, the reduced model in panel A for level of concern includes the interaction terms in addition to the consequence, type and order variables. The reduced models for relative level of substantive testing and number of audit hours in panels B, C and D include only the personal, firm or average consequence variable respectively, as an independent variable. None of the consequence variables in these models has a coefficient that is significantly greater than zero.

The coefficient for the average consequence variable in the reduced model for level of concern for PPE was not significantly greater than zero. The interactions between order and consequence level and between typeC and consequence have coefficients that are significantly greater than zero. The average consequence level, order of tasks, typeC, consequence by order interaction and consequence by typeC interaction are able to explain approximately 15 percent of the variation in concern levels for PPE.

\[ \text{A reduced model which included the three-way interaction between consequence, order and type (as well as all possible two-way interactions) resulted in an insignificant three-way interaction (p-value = .49).} \]
Table 12. Reduced Consequence Models for PPE - Results\textsuperscript{33}

Panel A: Reduced Model for Level of Concern - Average Consequence Assessment

\[ LC = \beta_0 + \beta_1 (Cons_i) + \beta_2 (Order_i) + \beta_3 (TypeC_i) + \beta_4 (Order_i \times Cons_i) + \beta_5 (TypeC_i \times Cons_i) + \epsilon_i \]

<table>
<thead>
<tr>
<th>( \beta )</th>
<th>( LC_{PPE} ) Coefficient (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 )</td>
<td>6.09*** (.00)</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>-.32 (.94)</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>-2.47 (.11)</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>-1.96 (.20)</td>
</tr>
<tr>
<td>( \beta_4 )</td>
<td>.60** (.02)</td>
</tr>
<tr>
<td>( \beta_5 )</td>
<td>.38* (.07)</td>
</tr>
</tbody>
</table>

Adjusted \( R^2 \) for the Model: .15

\textsuperscript{33} Reported p-values having a subscript 1 are for one-sided tests, other p-values are for two-sided tests.
Table 12 (continued)

Panel B: Reduced Models for Relative Substantive Testing and Audit Hours - Personal Consequence Assessment

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Coefficients (p-value)</th>
<th>R² for the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant</td>
<td>Personal Conseq.</td>
</tr>
<tr>
<td>LSTppE</td>
<td>6.41*** (.00)</td>
<td>.04 (.32)</td>
</tr>
<tr>
<td>AHppE</td>
<td>12.68*** (.00)</td>
<td>.38 (.23)</td>
</tr>
</tbody>
</table>

Panel C: Reduced Models for Relative Substantive Testing and Audit Hours - Firm Consequence Assessment

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Coefficients (p-value)</th>
<th>R² for the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant</td>
<td>Firm Conseq.</td>
</tr>
<tr>
<td>LSTppE</td>
<td>6.58*** (.00)</td>
<td>.00 (.48)</td>
</tr>
<tr>
<td>AHppE</td>
<td>11.33*** (.00)</td>
<td>.52 (.18)</td>
</tr>
</tbody>
</table>

Panel D: Reduced Models for Relative Substantive Testing and Audit Hours - Average Consequence Assessment

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Coefficients (p-value)</th>
<th>R² for the Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant</td>
<td>Average Conseq.</td>
</tr>
<tr>
<td>LSTppE</td>
<td>6.43*** (.00)</td>
<td>.03 (.38)</td>
</tr>
<tr>
<td>AHppE</td>
<td>11.04*** (.01)</td>
<td>.62 (.16)</td>
</tr>
</tbody>
</table>

*p significant at α = .10, **p significant at α = .05, ***p significant at α = .01
6.5.3 Summary of the Effect of Consequence

Overall, the assessed consequence levels seemed to impact subjects' audit planning judgments for AR as expected except when the number of audit hours was estimated. However, although the coefficients for consequence were statistically significant in single independent variable models, the percentage of variability in the dependent variable being explained is not very high ($R^2$ ranges from 5 percent to 17 percent). The second column of table 13 presents Pearson correlations between probability and consequence assessments for AR. The probability and personal consequences pair and the probability and average consequence level pair each have correlations of about 0.41. Probability and firm consequence level have a correlation of about 0.25. General linear tests of the reduced models including only probability and corresponding models that include probability and a consequence assessment as independent variables indicate that the consequence assessment does not provide a significant increase in explanatory power over the probability model (at an alpha level less than or equal to 5 percent). The firm consequence assessment did provide additional explanatory power for the level of concern for AR at an alpha level of 10%. Auditors may consider the expected loss to the firm when considering their relative concern about different accounts in an audit.

Consequence does not appear to provide significant explanation for planning judgments for PPE, except in one case. For someone who completes the CJM task first and/or considers consequence to be the most important factor, the average consequence level has a positive impact on the level of concern for PPE. The last column of table 13 shows the correlations between probability and consequence for PPE. The assessments of probability and consequence levels are not as highly correlated as the same assessments for AR.
It is not clear from these results that the auditors considered consequence levels at all in making substantive planning judgments. The ability of consequence levels to explain level of concern and substantive testing levels for AR may be due to their correlations with probability. It appears that subjects may not have treated the probability and consequence assessments independently, but tended to assess similar levels for both. The results indicate that probability and consequence levels are not as highly correlated for PPE, but consequence did not provide any significant explanation for planning judgments for PPE.

Table 13. Pearson Correlations between Probability and Consequence

<table>
<thead>
<tr>
<th>Pair of Variables</th>
<th>Correlation - AR</th>
<th>Correlation - PPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability and Personal Consequence</td>
<td>.41**</td>
<td>.20</td>
</tr>
<tr>
<td>Probability and Firm Consequence</td>
<td>.25**</td>
<td>.06</td>
</tr>
<tr>
<td>Probability and Average Consequence</td>
<td>.41**</td>
<td>.16</td>
</tr>
</tbody>
</table>

**significantly different from zero at $\alpha = .05$

34 The significance of the correlations was determined by comparing the appropriate test statistic to critical t-values for alpha levels equal to .10, .05 and .01. The test statistic used was $t/[(1-r^2)/(n-2)]^{1/2}$, where $r$ represents the Pearson correlation coefficient and $n$ represents the number of subjects. The number of subjects was 64 for each of the correlations.
6.6 Results of Tests of Hypothesis Three

Hypothesis three (H3) predicts that there will be an interaction between probability and consequence. Specifically, H3 predicts that the effect of a change in probability (consequence) on substantive testing judgments will be larger when the level of consequence (probability) is high. Based on this hypothesis, the coefficient for the probability by consequence interaction term, $\beta_3$, is expected to be significantly greater than zero. The results of the full models (table 8) indicate that there is not a significant positive interaction between assessed probability and assessed consequence. Therefore, H3 is not supported by the data collected in this study.

6.6.1 Reduced Models for the Probability by Consequence Interaction

Several reduced models were examined. Each model included a dependent variable and the probability variable, a consequence variable (personal, firm or average), and the interaction variable. None of these reduced models resulted in a significant interaction between probability and consequence. The specification of the reduced model and the results associated with the average consequence variable are reported in table 14.
Table 14. Reduced Regression Models for the Probability by Average Consequence Interaction - Results

\[ \text{LC}_{ij} (\text{LST}_{ij}, \text{AH}_{ij}) = \beta_0 + \beta_1 (\text{Prob}_{ij}) + \beta_2 (\text{Cons}_{ij}) + \beta_3 (\text{Prob}_{ij} \times \text{Cons}_{ij}) + \epsilon_{ij} \]

<table>
<thead>
<tr>
<th></th>
<th>LC\text{AR} Coeff. (p-value)</th>
<th>LST\text{AR} Coeff. (p-value)</th>
<th>AH\text{AR} Coeff. (p-value)</th>
<th>LC\text{PPE} Coeff. (p-value)</th>
<th>LST\text{PPE} Coeff. (p-value)</th>
<th>AH\text{PPE} Coeff. (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>.73 (.73)</td>
<td>3.98* (.06)</td>
<td>17.64 (.46)</td>
<td>1.85 (.21)</td>
<td>5.70*** (.00)</td>
<td>15.6* (.06)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>1.10*** (.00)_1</td>
<td>.64* (.06)_1</td>
<td>4.45 (.16)_1</td>
<td>.77** (.02)_1</td>
<td>.28 (.16)_1</td>
<td>-.87 (.66)_1</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>.65** (.03)_1</td>
<td>.05 (.44)_1</td>
<td>3.32 (.19)_1</td>
<td>.12 (.32)_1</td>
<td>-.08 (.67)_1</td>
<td>-1.2 (.81)_1</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>-.09 (.92)_1</td>
<td>.01 (.45)_1</td>
<td>-.49 (.77)_1</td>
<td>-.03 (.71)_1</td>
<td>.01 (.40)_1</td>
<td>.42 (.12)_1</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>.45</td>
<td>.54</td>
<td>.00</td>
<td>.38</td>
<td>.28</td>
<td>.14</td>
</tr>
</tbody>
</table>

*p significant at $\alpha = .10$, **p significant at $\alpha = .05$, ***p significant at $\alpha = .01$

Reported p-values having a subscript 1 are for one-sided tests, other p-values are for two-sided tests.
6.7 Results of Tests of Hypotheses Four and Five

Hypothesis four (H4) predicts that if a subject considers probability to be the most important factor his or her substantive planning judgments will be more positively influenced by the assessed probability of error than judgments of subjects who do not consider probability to be the most important factor. H4 implies that $\beta_6$ in the full model (table 8) should be significantly greater than zero. Since $\beta_6$ is not statistically significant in any of the models, H4 appears not to be supported by the data collected in this study. Because a significant relationship may be obscured by the inclusion of the many other variables in the full model, reduced models were also used to test H4 (section 6.7.1).

Hypothesis five (H5) predicts that if a subject considers consequence to be the most important factor his or her substantive planning judgments will be more positively influenced by the assessed consequence level than judgments of subjects who do not consider consequence to be the most important factor. H5 implies that $\beta_7$ in the full model (table 8) should be significantly greater than zero. $\beta_7$ is statistically significant in the LCPP model, but H5 is not supported in any of the other models.

As discussed in section 6.5.2 and illustrated in panel A of table 12, the coefficient for the typeC by consequence interaction in a reduced consequence model for LCPP was significantly greater than zero. This result implies that subjects who consider consequence to be the most important factor would have a lower level of concern for PPE at low levels of consequence but increase concern at a faster rate than other subjects as consequence increases. The typeC subjects appear to react to the consequence level as expected. The results of other reduced models for the typeC variable are discussed in section 6.7.2.
6.7.1 Reduced Models for the TypeP Variable

A reduced model for each of the dependent variables that included probability, typeP and the probability by typeP interaction was analyzed. H1 predicts that the coefficient ($\beta_1$ in the reduced model) for the probability variable will be significantly greater than zero. H4 predicts that the coefficient ($\beta_3$ in the reduced model) for the probability by typeP interaction will be significantly greater than zero. No prediction is made about the coefficient of the typeP variable ($\beta_2$ in the reduced model). If $\beta_2$ were significantly different from zero, that would indicate that a typeP person would plan a different level of substantive testing than a non-typeP person when there is a very small probability of error (close to zero). As illustrated in table 3 (Chapter III), the score values for combinations of the three factors do not differ between the two prototypes for lowest valued or highest valued combinations.

Table 15 shows that being a typeP subject does not significantly affect substantive testing judgments in the expected direction. In the relative level of substantive testing model for PPE, the p-value for the typeP by probability interaction is .95, and the coefficient is negative.
Table 15. Reduced Regression Models for TypeP - Results\(^{36}\)

\[ LC_{ij} (LST_{ij}, AH_{ij}) = \beta_0 + \beta_1 (Prob_{ij}) + \beta_2 (TypeP_i) + \beta_3 (TypeP_i \times Prob_{ij}) + \epsilon_{ij} \]

<table>
<thead>
<tr>
<th></th>
<th>(LC_{AR}) Coeff. (p-value)</th>
<th>(LST_{AR}) Coeff. (p-value)</th>
<th>(AH_{AR}) Coeff. (p-value)</th>
<th>(LC_{PPE}) Coeff. (p-value)</th>
<th>(LST_{PPE}) Coeff. (p-value)</th>
<th>(AH_{PPE}) Coeff. (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\beta_0)</td>
<td>3.97*** (.00)</td>
<td>4.61*** (.00)</td>
<td>31.15*** (.00)</td>
<td>2.46*** (.00)</td>
<td>5.05*** (.00)</td>
<td>6.93** (.02)</td>
</tr>
<tr>
<td>(\beta_1)</td>
<td>.69*** (.00)(^1)</td>
<td>.63*** (.00)(^1)</td>
<td>2.37** (.03)(^1)</td>
<td>.66*** (.00)(^1)</td>
<td>.45*** (.00)(^1)</td>
<td>2.27*** (.00)(^1)</td>
</tr>
<tr>
<td>(\beta_2)</td>
<td>1.40 (.14)</td>
<td>-1.17 (.20)</td>
<td>16.76 (.11)</td>
<td>.10 (.90)</td>
<td>.42 (.46)</td>
<td>3.38 (.44)</td>
</tr>
<tr>
<td>(\beta_3)</td>
<td>-.23 (.90)(^1)</td>
<td>.21 (.11)(^1)</td>
<td>-2.40 (.90)(^1)</td>
<td>-.19 (.86)(^1)</td>
<td>-.22 (.95)(^1)</td>
<td>-1.50 (.93)(^1)</td>
</tr>
<tr>
<td>Adjusted (R^2)</td>
<td>.43</td>
<td>.55</td>
<td>.04</td>
<td>.42</td>
<td>.34</td>
<td>.16</td>
</tr>
</tbody>
</table>

**p significant at \(\alpha = .05\), ***p significant at \(\alpha = .01\)

\(^{36}\) Reported p-values having a subscript \(^1\) are for one-sided tests, other p-values are for two-sided tests.
6.7.2 Reduced Models for the TypeC Variable

A reduced model for each of the dependent variables that included a consequence, typeC and the consequence by typeC interaction was analyzed. H2 predicts that the coefficient (β1 in the reduced model) for the consequence variable will be significantly greater than zero. H5 predicts that the coefficient (β3 in the reduced model) for the consequence by typeC interaction will be significantly greater than zero. No prediction is made about the coefficient of the typeC variable (β2 in the reduced model). If the variable were significantly different from zero, that would indicate that a typeC person would plan a different level of substantive testing than a non-typeC person when there is a very small consequence associated with not detecting a material error (close to zero). For the same reason as stated for the typeP variable, the typeC variable is not expected to significantly differ from zero.

Panel A of table 16 shows that being a typeC subject significantly affects the change in the level of concern for AR in the expected direction when considering the firm consequence level (p-value = .09). The reduced model for level of concern implies that someone who is not a typeC would rate his or her level of concern at approximately 6.6, regardless of the level of firm consequence since the coefficient for the firm consequence variable (β2 in the reduced model) is not significantly different from zero. A typeC subject, on the other hand, would start at the same concern level but is expected to increase his or her concern as the level of firm consequences increases. For all three dependent variables, reduced models that include the personal or average consequence level result in coefficients for the typeC by consequence interaction that are not significantly different from zero.
Table 16. Reduced Regression Models for TypeC - Results37

\[ LC_{ij} (LST_{ij}, AH_{ij}) = \beta_0 + \beta_1 (Cons_{ij}) + \beta_2 (TypeC_i) + \beta_3 (TypeC_i \times Cons_{ij}) + \epsilon_{ij} \]

Panel A: Reduced Models Including Firm Consequence and TypeC for AR

<table>
<thead>
<tr>
<th></th>
<th>( L_{CAR} ) Coefficient (p-value)</th>
<th>( LST_{CAR} ) Coefficient (p-value)</th>
<th>( AH_{CAR} ) Coefficient (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 )</td>
<td>6.63*** (.00)</td>
<td>6.51*** (.00)</td>
<td>43.87*** (.00)</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>.14 (.23)</td>
<td>.17 (.22)</td>
<td>.44 (.40)</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>-2.71 (.17)</td>
<td>-1.25 (.56)</td>
<td>-8.22 (.63)</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>.36* (.09)</td>
<td>.19 (.26)</td>
<td>.44 (.43)</td>
</tr>
<tr>
<td>Adjusted R(^2) for the Model</td>
<td>.09</td>
<td>.01</td>
<td>.00</td>
</tr>
</tbody>
</table>

*p significant at \( \alpha = .10 \), ***p significant at \( \alpha = .01 \)

37 Reported p-values having a subscript 1 are for one-sided tests, other p-values are for two-sided tests.
Panel B: Reduced Models Including Average Consequence and TypeC for PPE

Table 16 (continued)

<table>
<thead>
<tr>
<th></th>
<th>LST&lt;sub&gt;PPE&lt;/sub&gt; Coefficient (p-value)</th>
<th>AH&lt;sub&gt;PPE&lt;/sub&gt; Coefficient (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>β&lt;sub&gt;0&lt;/sub&gt;</td>
<td>6.59*** (.00)</td>
<td>14.21** (.01)</td>
</tr>
<tr>
<td>β&lt;sub&gt;1&lt;/sub&gt;</td>
<td>-.03 (.59)&lt;sub&gt;1&lt;/sub&gt;</td>
<td>-.18 (.58)&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>β&lt;sub&gt;2&lt;/sub&gt;</td>
<td>-.46 (.68)</td>
<td>-8.00 (.28)</td>
</tr>
<tr>
<td>β&lt;sub&gt;3&lt;/sub&gt;</td>
<td>.16 (.20)&lt;sub&gt;1&lt;/sub&gt;</td>
<td>2.07** (.05)&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

Adjusted R<sup>2</sup> for the Model

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.00</td>
</tr>
</tbody>
</table>

**p significant at α = .05, ***p significant at α = .01
Panel B of table 16 indicates that being a typeC subject significantly affects the audit hours planned for PPE. According to the results of the AH model in panel B, all subjects are expected to plan about 14 hours for PPE when average consequence level is very low. Only consequence types are expected to increase audit hours as consequence level increases. Other subjects would be expected to maintain a similar number of audit hours regardless of changes in consequence level.

6.8 Results of Tests of Hypothesis Six

Hypothesis six, H6, predicted that using the alternative dependent variables (LC and LST) would result in a clearer and stronger relationship between the probability of error in an account and the level of substantive testing planned for the account. It appears that probability of error was better able to explain alternative dependent variables than it explained AH. The full regression models in table 8 show that the probability variable is significant at an alpha level less than or equal to ten percent for the LC<sub>AR</sub>, LST<sub>AR</sub>, and LC<sub>PPE</sub>. In regression models where probability of error is the only independent variable (see tables 9 and 10 in sections 6.4.1 and 6.4.2), the R<sup>2</sup> is higher and the p-value for probability is lower when LC and LST are used as dependent variables compared to when AH is used. R<sup>2</sup> (the percent of variation in the planning judgment explained by variation in probability) is 44 percent for the LC<sub>AR</sub>, 56 percent for LST<sub>AR</sub> and four percent for AH<sub>AR</sub>. The reduced probability models for PPE also include the order variable and so they are not directly comparable to the reduced probability models for AR.

---

38 Results of reduced models including personal or firm consequence levels are qualitatively similar to the reduced models that include the average consequence levels.

39 The reduced models using LST (level of substantive testing) and AH (audit hours) for PPE shown in Table 10 included the order variable and the interaction between probability and order variable since these were also significant.
The Pearson correlations for probability and LC, probability and LST and probability and AH are 0.66, 0.75 and 0.19 percent respectively for AR. The correlations between each of the alternative dependent variables and probability for AR are significantly different from zero at an alpha level less than .01. The correlation between audit hours and probability of error for AR is not significantly different from zero at an alpha level of .10. The Pearson correlations between the assessed probability and each of the dependent variables for PPE are 0.64 for probability and the level of concern, 0.56 for probability and the relative level of substantive testing and about 0.39 for probability and the number of audit hours. Although the correlations are larger for the alternative dependent variables, all three of the correlations are significantly different from zero at an alpha level less than .01.\(^{40}\) Thus, H6 appears to be strongly supported by the data for AR, but not as strongly for PPE.

The correlations between the dependent variables provide some insight into why there is a difference in the extent of support for H6 between the two accounts. Table 17 reports these correlations. The correlation between the alternative dependent variables and the audit hours planned for PPE is fairly high (about 0.60). The correlation between the alternative dependent variables and audit hours planned for AR is not nearly as high (0.29 and 0.34). The subjects’ planning judgments for AR appear to have different amounts of shared variability depending on the judgments being compared. The subjects’ planning judgments for PPE on the other hand have a high level of shared variability.

\(^{40}\) Significance of these correlations was computed in the same way as for the correlations reported in table 13.
Table 17. Pearson Correlations between Dependent Variables

<table>
<thead>
<tr>
<th>Pair of Variables</th>
<th>Correlation - AR</th>
<th>Correlation - PPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Concern and Relative Level of Testing</td>
<td>.63***</td>
<td>.64***</td>
</tr>
<tr>
<td>Relative Level of Testing and Audit Hours</td>
<td>.34***</td>
<td>.60***</td>
</tr>
<tr>
<td>Level of Concern and Audit Hours</td>
<td>.29**</td>
<td>.60***</td>
</tr>
</tbody>
</table>

**significant at $\alpha = .05$, ***significant at $\alpha = .01^{41}$

---

41 Significance of these correlations was computed in the same way as for the correlations reported in table 13. The number of subjects was only 60, however, in correlations that included audit hours for PPE.
6.9 Importance of Factors and Trade-offs Made between Factors

Table 18 indicates the factor or factors that the auditors considered to be the most important when performing the CJM task.

Table 18. Most Important Factor

<table>
<thead>
<tr>
<th>Probability</th>
<th>Effort</th>
<th>Consequence</th>
<th>Probability and Consequence</th>
<th>No Clear Preference</th>
<th>Total Auditors</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>0</td>
<td>34</td>
<td>3</td>
<td>2</td>
<td>65</td>
</tr>
</tbody>
</table>

The major emphasis in analyzing the CJM rankings was on the trade-offs made between pairs of factors. 42 Twenty-five subjects made only trade-off violations of joint independence. The most common trade-offs made were between probability and effort (approximately 43 percent of the trade-offs made). When consequence levels were high, probability (effort) was more important in 60 percent (40 percent) of these cases. As the consequence level decreased, effort (probability) became more important. Approximately 32 percent of the trade-offs were between effort and consequence. When probability was high, about half (52 percent) of these cases ranked effort as being more important with consequence becoming more important as probability decreased.

---

42 Although it was not expected, nineteen subjects had simple independence violations in their rankings. The most common simple independence violations, 53 percent of all of these types of violations, involved the effort factor. Twenty-eight percent of the violations involved the probability factor and 19 percent involved the consequence factor. Twenty-one subjects did not make any independence violations at all when ordering the 27 cards. Their preference orderings could be predicted by an additive model (as described in Appendix F). Nine of these subjects indicated that probability was the most important factor and 12 indicated that consequence was the most important factor of the three.
Probability and consequence trade-offs were the least frequent comprising about 25 percent of all trade-offs made. When effort was low, 60 percent (40 percent) of these situations ranked consequence (probability) more important with probability (consequence) increasing in importance as the effort level increased.
CHAPTER VII
DISCUSSION

As described in Chapter I, this research was motivated by witnessing an apparent lack of significant relationship between the probability of error in an account and the number of audit hours planned for the account. The major objective of the study was to explore potential explanations for the apparent lack of association. The first section of this chapter discusses the results that relate to the major objective and additional research that may complement or expand upon the results found here. Subsequent sections of this chapter discuss the additional contributions expected from this research as presented in Chapter I. The final section of this chapter summarizes the implications of this study for audit research and audit firms.

7.1 Results Associated with the Major Objective of this Study

The major objective of this research was to explore potential explanations for the lack of significant association between probability of error and planned audit hours in previous audit planning research. The results of this study were consistent with the previous research in that probability of error did not provide significant explanation for planned audit hours for AR. The most plausible explanation appears to be that a large amount of the variation in planned audit hours between subjects is not variability that is shared with subjects' assessed probabilities of error. Variation in alternative dependent variables such as relative level of concern and relative level of substantive testing was explained by probability assessments for AR. Variability in audit hours may be better
explained by differences in audit planning specifics between subjects that lead to different numbers of audit hours. As suggested in figure 1, types of tests, extents of tests and levels of audit personnel assigned to perform the tests may account for much of the unexplained variation in audit hours.

The other two explanations proposed were not supported by the data in this study. Neither of the consequence assessments (personal or firm) provided significant explanation for variation in audit hours planned for AR. The subject's type did not appear to enhance his or her consideration of probability or consequence in the AR audit hour models. Thus, the subject's most important factor is probably not the individual difference that would explain the large variation in audit hours.

7.1.1 Implications of Results Associated with the Major Objective

The results of this study suggest that probability of error in AR is associated with audit planning judgments. The large amount of unexplained variation in audit hours witnessed in this experiment and others may not be a problem. The audit environment has at least two characteristics that may lead to less variation in planned audit hours than what appears in experimental settings. First, within an audit firm an individual auditor is typically not fully responsible for determining the number of audit hours planned. An audit team consisting of several auditors with various experience levels would ultimately determine the number of audit hours on an engagement. If the final audit plan is the result of a group consensus, variation in audit hours may be eliminated through group discussion. It is not clear that the high individual variability in estimating audit hours would persist in an actual audit setting.

Second, the variability introduced by individual auditors' choices of tests, sample sizes and level of personnel to perform testing may be mitigated by the amount of information available in an audit that is not available in experimental approaches to
studying audit planning. For an actual audit, the auditor would have much more information about the client company, knowledge of the client's management and experience with and knowledge of the industry. It may be unrealistic to expect planned audit hours in an experiment to be significantly associated with the probability of error. In an experimental setting individual auditors' judgments about relative levels of substantive testing appear to be more predictable than audit hours. The use of relative substantive testing levels may be more appropriate because it does not require subjects to make the intermediate decisions (types of tests, sample sizes and level of personnel) that are required to estimate audit hours.

Future audit planning research may include relative testing levels as a dependent variable to determine whether results would be consistent with this study. In addition, the typical experimental research approach may be expanded. After asking individuals to indicate relative testing levels groups could be formed and asked to determine the number of audit hours they would plan for AR. This step would indicate if high variability in the number of audit hours planned would persist across groups.

7.2 Link between Theoretical and Experimental Audit Planning Research

One of the expected contributions of this research was to establish a link between the theoretical audit planning research and the experimental approach to studying audit planning. This study explicitly included consequence as a variable for the auditors to assess. Contrary to expectations based on theoretical research, consequence did not generally provide significant explanation for audit planning judgments beyond what could be explained by the probability assessments. The one exception was firm consequence which provided significant additional explanation (at an alpha level of .10) for the level of concern for AR.
When subjects made planning judgments for AR, consequence levels were highly correlated with the probability of error in the account. This phenomenon, witnessed in the SAC task, makes it difficult to determine whether subjects were able to distinguish between probability and consequence as distinct conceptual factors that should impact their judgments. In the CJM task most subjects clearly indicated a reliance on either probability or consequence as the most important factor. Very few subjects (three out of 65) indicated that both probability and consequence were most important.

Future research may be aimed at incorporating the concept of consequence as distinct from the probability of error in an audit case similar to what was used in the SAC task. The manner in which the consequence levels were manipulated in this study may not have been sufficient to test the true effect of the consideration of consequence.

In addition, future research may use managers and partners as subjects. The significance of the firm consequence assessment for the level of concern in the current study suggests that the expected future costs to the firm were considered by some auditors. With auditors who are closer to the partner level, it is expected that the possibility of a lawsuit would be more salient and perhaps have a greater influence on audit judgments than it did for the auditors in this study.

7.3 Experimental Evidence on More than One Account within an Audit

The inclusion of two accounts in this study revealed differences in the planning judgments made across accounts. For AR, auditors' relative testing levels were highly correlated with the assessed probability of error in the account. In addition, auditors appeared to positively relate consequence levels to probability levels. Consequence levels on their own did not provide significant explanation for planning judgments above what could be explained by probability, except when considering the level of concern for AR.
For PPE, probability was positively associated with relative levels of concern and relative substantive testing levels. The positive impact of assessed probability on substantive testing levels was enhanced for subjects who completed the CJM task first. Assessed probability of error was significantly related to audit hours for subjects who completed the CJM task first. These subjects tended to plan lower levels of substantive testing and fewer audit hours for low levels of probability but increased testing and hours more as probability increased compared to other subjects.

Subjects' consequence assessments for PPE were not highly correlated with their probability assessments for PPE. Consequence did not provide significant explanation for planning judgments except in two cases. First, subjects who completed the CJM task and considered consequence to be the most important factor increased their level of concern for PPE as consequence levels increased. Second, subjects who considered consequence to be the most important factor increased audit hours for PPE as consequence levels increased.

Overall, these results suggest that auditors may be better able or more willing to consider the probability of error when indicating concern or relative testing levels for AR. The complexity of auditing AR, compared to PPE, may be the reason for audit hours not being associated with probability or consequence for AR but associated with probability and consequence for some subjects for PPE.

Future research may also include more than one account to determine if similar differences between accounts would be witnessed. A survey instrument to managers and partners of audit firms may explore whether they think the accounts should be treated differently.
7.4 Evidence about Auditors’ Priorities

The results of the CJM task indicated that forty percent of the subjects considered probability to be the most important of the three factors and about fifty-two percent considered consequence to be the most important factor. Although the auditors’ types did not influence their planning judgments in the SAC task, it is possible that these differences in emphasis would have an impact on other decisions made throughout the audit. For example, as the audit progresses compromises may be made and less work may be performed on a low consequence account even if the probability of error is high because the subject considers consequence to be the most important factor.

Several participants made trade-offs between factors in the CJM task. The most common trade-offs were between probability and effort. When forming a preference for which types of accounts would receive high levels of testing, fifteen auditors emphasized probability when the consequence level was high, but emphasized effort when the consequence level was low.

Future research that investigates the potential effect of different importance beliefs on specific critical areas of the audit process may be beneficial to audit firms as well as audit researchers.

The trade-offs described above imply that some auditors may attempt to minimize the expected cost of the audit. Research on auditors’ utility functions would help to shed light on what auditors value and what they try to optimize in planning and performing the audit.
7.5 Conclusion

7.5.1 Implications for Audit Planning Research

The results of this study appear to provide a plausible explanation for the lack of a significant relationship between probability of error and audit hours in Bedard and Wright (1992). For future experimental audit planning research, researchers may consider including an alternative dependent variable in lieu of or in addition to the number of audit hours.

This study provides evidence that, as professed by previous researchers, probability of error in an account does influence audit planning judgments. The findings presented here may provide some direction for future researchers to continue to sort out the various influences on the audit planning process. The results provide inconsistent evidence between the two tasks about the auditors' consideration of consequences in making planning judgments. A more effective research design is needed to simulate the auditors' incorporation of expected future costs in the audit planning process.
7.5.2 Implications for Audit Practice

Although this is only one study with relatively few (65) auditors represented, the results may be of interest to auditing firms. First, the subjects in this study appeared to be about evenly split between those who thought probability was most important and those who thought consequence was most important. If the partners in auditing firms are more consistent in their consideration of relative importance of the three factors, they may be concerned that auditors of lower ranks do not adhere to the same priorities. This would appear to be especially worrisome if an auditor's type (most important factor) is found to have significant influence on audit planning judgments.

The reaction of auditors' relative testing levels to probability of error would seem to be encouraging to audit firms. However, the large variability in the number of audit hours associated with each planned level of testing may be disturbing. Each firm may consider its own audit planning process and determine whether the diversity in the number of audit hours planned should be of concern to them, in particular.
Low (P1): Overall, the control environment has been assessed as effective. This account is one that is affected by material error with relatively low frequency in the client’s industry. The accounting system appears to be reliable (from previous audit). The data process is routine and small amounts are involved in individual transactions. Accounting personnel are highly qualified and there is low turnover of personnel. Preliminary analytical review procedures involving this account did not yield any unexpected results.

Medium (P2): Overall, the control environment has been assessed as effective. This account is one that is affected by material error with relatively moderate frequency in the client’s industry. The account involves an accounting estimate that is material in dollar amount. The estimation procedure is well documented and management’s assumptions are clear and seem reasonable. Preliminary analytical review procedures involving this account show a difference from what was expected.

High (P3): Overall, the control environment has been assessed as effective. This account is one that is affected by material error with relatively moderate frequency in the client’s industry. The account involves an accounting estimate that is material in dollar amount. The estimation procedure is not well documented and is subject to extensive management discretion. Management’s assumptions providing the basis of the estimate are not clear. Preliminary analytical review procedures involving this account indicate a significant difference from what was expected.
Figure 6 (continued)

EFFORT

**High (E1):** If this account were classified as a high audit emphasis account, there would be several additional procedures prescribed. In addition, the scope of two of the procedures required for basic testing would be significantly increased. The additional effort required to increase substantive testing beyond the minimal level would be substantial.

**Medium (E2):** If this account were classified as a high audit emphasis account, there would be one additional procedure prescribed. In addition, the scope of two of the procedures required for basic testing would be moderately increased. The additional effort required to increase substantive testing beyond the minimal level would be noticeable but not substantial.

**Low (E3):** If this account were classified as a high audit emphasis account, there would be no additional procedures prescribed. The scope of one procedure required for basic testing would be moderately increased. The additional effort required to increase substantive testing beyond the minimal level would hardly be noticeable.

CONSEQUENCE

**Low (C1):** Misstatement of this account will not have any impact on debt covenant provisions. There is only a remote possibility that misstatement of this account would cause the financial statements not to be presented fairly.

**Medium (C2):** Misstatement of this account will not have any impact on debt covenant provisions. There is a possibility that misstatement of this account would cause the financial statements not to be presented fairly.

**High (C3):** Misstatement of this account would significantly impact debt covenant provisions. There is a good possibility that misstatement of this account would cause the financial statements not to be presented fairly.
P3: Overall, the control environment has been assessed as effective. This account is one that is affected by material error with relatively moderate frequency in the client's industry. The account involves an accounting estimate that is material in dollar amount. The estimation procedure is not well documented and is subject to extensive management discretion. Management's assumptions providing the basis for the estimate are not clear. Preliminary analytical review procedures involving this account indicate a significant difference from what was expected.

E1: If this account were classified as a high audit emphasis account, there would be several additional procedures prescribed. In addition, the scope of two of the procedures required for basic testing would be significantly increased. The additional effort required to increase substantive testing beyond the minimal level would be substantial.

C2: Misstatement of this account will not have any impact on debt covenant provisions. There is a possibility that misstatement of this account would cause the financial statements not to be presented fairly.

Figure 7. Sample of Card Used in CJM Task
Description of Exercise and Instructions

As you know from your experience, there are many variables and much uncertainty associated with planning an efficient and effective audit. For the purposes of this research, some of the variables you might consider are classified into three categories, or factors, called "Probability," "Effort," and "Consequence." "Probability" refers to the likelihood that there is a material error in an account. "Effort" refers to the additional audit effort required to increase substantive testing of the account from a minimal level to a high level. "Consequence" refers to the expected cost to you and the audit firm if a material error is not detected during the audit but is subsequently discovered after the audit report is issued.

To gain some insight into how probability of a material error, effort associated with increasing testing levels, and consequences might influence levels of substantive testing planned for various accounts, you are asked to do a preference ordering exercise. You are given 27 cards in this envelope, each one describing the characteristics of an account using the three factors defined above. On each of these cards, you will first see a description which is intended to convey a low, medium or high level of probability of a material error existing in the account. The description that represents the low level of probability is labeled as "P1." The description that represents the medium level of probability is labeled as "P2." The description that represents the high level of probability is labeled as "P3." The second item on each card is a description which is
intended to convey a high, medium or low level of effort required to increase the level of substantive testing from a minimal level to an extensive level. The description that represents the high level of effort is labeled as “E1.” The description that represents the medium level of effort is labeled as “E2.” The description that represents the low level of effort is labeled as “E3.” The third item on each card is a description which is intended to convey a low, medium or high level of consequence expected if there is a material error in an account and it is not detected during the audit but is subsequently discovered. The description that represents the low level of consequence is labeled as “C1.” The description that represents the medium level of consequence is labeled as “C2.” The description that represents the high level of consequence is labeled as “C3.”

You are asked to consider each card as representing an account. You may think of a particular account such as "cash," or "accounts receivable" when you see one of the cards but no account names are written on the cards and they are not intended to represent certain named accounts. For each "account" please determine how you would rank it compared to the other "accounts" (represented by cards) in terms of preferring to perform a high level of substantive testing on the account. For example, the card labeled #1 indicates a high probability of material error, a low level of effort required to increase substantive testing and a high level of consequences (high expected future cost) associated with not detecting an existing material error in the account (P3, E3, C3). Compared to any other combination of probability, effort and consequence, an account with these characteristics (as described on the card labeled #1) would be the account that you would most prefer to extend substantive testing beyond a minimal level. There is also a card labeled #27, containing a low probability of material error, a high effort required to extend testing and a low level of consequences (P1, E1, C1), that represents an account that you would least prefer to extend substantive testing beyond some minimal level.
Using the cards prenumbered as 1 and 27 as endpoints, please put the remaining 25 cards in order according to your preference for performing a high level of substantive testing. For the purpose of this exercise, assume that you have decided not to rely on the internal control systems associated with these accounts due to the high expense of testing the controls.

To help you order the cards, you may place card #1 to one side and card #27 to the other side with the remaining 25 cards in a pile in front of you. (The 25 cards are in random order.) Take the top card from the pile and place it between #1 and #27. Take the second card from the pile and place it in order between #1 and the first card you placed or between the first card you placed and #27. Continue taking one card at a time from the pile of cards and placing it in order between the cards you have already laid out in front of you. After taking the last card from the pile you should have 27 cards in front of you ordered from #1 to #27 in the order that you would prefer to increase substantive testing on the described "account" beyond the minimal level. Please check your sequence by comparing adjacent cards. Verify that each card represents an account on which you would prefer to place more audit emphasis compared to the card that comes right after it. After ensuring that the cards as ordered represent your preferences (with #1 on top and #27 on the bottom), put a check mark on the cover card, and place a rubber band around the ordered stack of cards. After completing the preference ordering exercise, please place this document and the cards back in the envelope.

Thank you.
APPENDIX C
VERSION A OF SAC TASK

Introduction

This case deals with planning substantive tests for accounts receivable and property, plant and equipment. Presented below are a description of the audit client, general information about the client's control environment, and information about each of the two accounts. Comparative financial statements and schedules for each of the accounts are provided. After reading the information and looking at the financial statements and schedules, please provide the responses requested on pages 13 - 18 of this packet. Feel free to make any notes, comments, calculations, etc. directly on these pages. Even though you may not have all the information you would have on an actual audit engagement, try to imagine yourself in an actual engagement and express your preferences as best you can. There are no right or wrong answers to this case. Your honest and thoughtful responses will be greatly appreciated. Your responses will remain confidential.

JACOBS MANUFACTURING COMPANY

Jacobs Manufacturing Company is a medium-sized manufacturing entity whose major product line consists of submersible electric motors. Customers use Jacobs motors to pump water from deep below the earth's surface and for oil wells. The company also produces fractional horsepower electric motors. Jacobs Manufacturing has been in business for 20 years.

The company's executive offices and manufacturing plant are located in Indiana. Approximately 800 people are employed by the company.

Jacobs has approximately 600 shareholders. The company's stock is traded on the NASDAQ national market. The Chairman of the Board and Chief Executive Officer, William
Jacobs, owns 5% of the outstanding common stock. No other officer or director owns more than 1%.

The Company's Board of Directors consists of the Chairman and six outside directors. In addition to Mr. Jacobs, other officers are: David Coleman, Vice President of Operations; Shirley Johnson, Vice President of Marketing; Roger Hopkins, Vice President of Engineering; Howard McDonald, Vice President of Finance and Lewis Neville, Secretary and Treasurer.

By taking advantage of niche opportunities, Jacobs has maintained its market share of about 20% over the past several years. Their strategy is to become an integral part of some customers' engineering and manufacturing capabilities. They have developed long-term mutually beneficial relationships with many of their customers.

Operating Activities. The Company is subject to the normal reporting requirements of a publicly-owned manufacturing concern. The company's accounting system is computerized. A preliminary evaluation of the internal control system has been done, but due to the expected high cost of testing the controls, it has been decided that the internal control system will not be relied on. Preliminary financial statements for 1992 and audited financial statements for 1991 and 1990 are presented on pages 3 and 4.

Industry. The electric motor market is highly competitive with both large and small suppliers vying for market share. Suppliers sell electric motors to original equipment manufacturers of pumps, petroleum pumping equipment, compressors, fans, heating and air conditioning equipment, swimming pool equipment, medical furniture and business machines. Motors are also sold in the replacement market through independent distributors and repair shops.

Industry sales are expected to grow steadily over the next 5 years (4 - 6% each year).
JACOBS MANUFACTURING COMPANY


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Net Sales</td>
<td>$103,131</td>
<td>$92,034</td>
<td>$88,493</td>
<td>12%</td>
<td>4%</td>
</tr>
<tr>
<td>% of Sales</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost of Goods Sold</td>
<td>77,142</td>
<td>69,578</td>
<td>67,166</td>
<td>11%</td>
<td>4%</td>
</tr>
<tr>
<td>Gross Margin</td>
<td>$25,989</td>
<td>$22,456</td>
<td>$21,327</td>
<td>16%</td>
<td>5%</td>
</tr>
<tr>
<td>% of Sales</td>
<td>25.2</td>
<td>24.4</td>
<td>24.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selling, General and Administrative Expenses</td>
<td>13,486</td>
<td>11,819</td>
<td>11,271</td>
<td>14%</td>
<td>5%</td>
</tr>
<tr>
<td>Other Operating Expenses</td>
<td>1,072</td>
<td>1,002</td>
<td>938</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>Operating Income</td>
<td>$11,431</td>
<td>$9,635</td>
<td>$9,118</td>
<td>19%</td>
<td>6%</td>
</tr>
<tr>
<td>Interest Expense</td>
<td>629</td>
<td>639</td>
<td>683</td>
<td>(2)%</td>
<td>(6)%</td>
</tr>
<tr>
<td>Income before Income Taxes</td>
<td>$10,802</td>
<td>$8,996</td>
<td>$8,435</td>
<td>20%</td>
<td>7%</td>
</tr>
</tbody>
</table>

126
JACOBS MANUFACTURING COMPANY


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash and Marketable Securities</td>
<td>$3,779</td>
<td>$5,454</td>
<td>$4,447</td>
<td>(31)%</td>
<td>23%</td>
</tr>
<tr>
<td>Receivables</td>
<td>$10,143</td>
<td>$7,446</td>
<td>$7,197</td>
<td>36</td>
<td>3</td>
</tr>
<tr>
<td>Less: Allow. for Doubtful Accounts</td>
<td>$(846)</td>
<td>$(720)</td>
<td>$(702)</td>
<td>16 (allow.)</td>
<td>4 (allow.)</td>
</tr>
<tr>
<td>Inventories</td>
<td>7,714</td>
<td>7,709</td>
<td>7,366</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Other Current Assets</td>
<td>1,810</td>
<td>1,858</td>
<td>1,908</td>
<td>(3)</td>
<td>(3)</td>
</tr>
<tr>
<td>Current Assets</td>
<td>$22,600</td>
<td>$21,737</td>
<td>$20,216</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Property, Plant and Equipment (net)</td>
<td>13,046</td>
<td>11,192</td>
<td>9,663</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>Other Assets</td>
<td>10,503</td>
<td>10,109</td>
<td>10,381</td>
<td>4</td>
<td>(3)</td>
</tr>
<tr>
<td>Total Assets</td>
<td>$46,149</td>
<td>$43,038</td>
<td>$40,260</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Notes Payable to Banks</td>
<td>$62</td>
<td>$53</td>
<td>$56</td>
<td>17%</td>
<td>(5)%</td>
</tr>
<tr>
<td>Current Maturities -- Long-term Debt</td>
<td>453</td>
<td>445</td>
<td>433</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Accounts Payable</td>
<td>3,001</td>
<td>2,722</td>
<td>2,523</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Accrued Expenses</td>
<td>9,868</td>
<td>7,850</td>
<td>7,356</td>
<td>26</td>
<td>7</td>
</tr>
<tr>
<td>Other Current Liabilities</td>
<td>321</td>
<td>440</td>
<td>394</td>
<td>(27)</td>
<td>12</td>
</tr>
<tr>
<td>Current Liabilities</td>
<td>$13,705</td>
<td>$11,510</td>
<td>$10,762</td>
<td>19</td>
<td>7</td>
</tr>
<tr>
<td>Long-term Debt</td>
<td>7,287</td>
<td>7,770</td>
<td>7,331</td>
<td>(6)</td>
<td>6</td>
</tr>
<tr>
<td>Other Long-term Liabilities</td>
<td>3,345</td>
<td>3,599</td>
<td>3,667</td>
<td>(7)</td>
<td>(2)</td>
</tr>
<tr>
<td>Total Liabilities</td>
<td>$24,337</td>
<td>$22,879</td>
<td>$21,760</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Stockholders' Equity</td>
<td>$21,812</td>
<td>$20,159</td>
<td>$18,500</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Total Liabilities and Stockholders' Equity</td>
<td>$46,149</td>
<td>$43,038</td>
<td>$40,260</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>
Other Information

1. Control Environment and Management

The control environment at Jacobs is adequate. Control mechanisms are in place throughout the operations of the company. The president is very involved in the technical aspects of the business and leaves operational duties to the other officers and their managers. Mr. Jacobs is concerned, however, about maintaining market share in the future in light of a poor economy. At the beginning of 1992, he implemented a new incentive bonus plan for sales managers. From talking to some of the managers, you sense great enthusiasm for the program. Managers receive a bonus for achieving gross sales above specified levels. The average bonus per sales manager in 1992 was $18,000, which was a significant portion of each manager's total compensation.

The company has been a client of your firm for three years. Although you have no other evidence that the chief accounting officer, Mr. Neville, is of low integrity, there have been minor disagreements with him in the past about accounting issues. He tends to exhibit a bias when applying accounting principles, choosing presentations or methods of estimation that show the company in the best light. A few audit adjustments were made in each of the past three years, but none involved accounts receivable or property, plant and equipment.

2. Sales Function

The sales department consists of 40 sales managers and their clerical staff. These managers receive periodic technical training to ensure thorough knowledge of the products.

Sales orders are typically taken over the telephone. The order may be customer initiated or solicited by the sales manager. Because of the customized nature of many of the pumps and motors, the sales manager may have to visit the customer's facility. After ascertaining the customer's needs, the manager enters order information into the computer, on a prenumbered sales order form. The sales order includes customer information, a general description of the product and an exact price, if available (from an annually updated price list) or an estimated price for custom orders. For new orders under $500, orders that would not increase a customer's balance to more than $500, and orders from repeat customers (within their preapproved credit limits), the sales manager indicates on the sales order, "credit O. K."
and electronically sends the sales order to the general accounting department. An order that requires an initial credit approval or an increase in the credit limit is sent directly to the credit manager. If approved by the credit department, the sales order is then sent to the general accounting department. If not approved, the sales order is returned to the sales manager, who contacts the customer.

The general accounting department, upon receiving the sales order, checks the order for accuracy and completeness. Copies of the approved sales order are then sent to the sales department, factory order department, shipping department and customer (as a verification of the order).

When the ordered items are completed, the factory sends the order to the shipping department where the items are compared to the sales order. A prenumbered shipping document is prepared. One copy is packed with the order, two copies are sent with the driver of the delivery truck and one copy is sent to the general accounting department. A prenumbered sales invoice is prepared in general accounting after comparing the shipping document to the original sales order. At that time, the sale and receivable are recorded and the sales invoice is sent to the customer.

3. **Accounts Receivable**

Most of the company's products are sold on normal trade terms, net 30, at published prices. Price lists are updated annually as needed. For these products the price is based on total cost to manufacture plus a markup. Some products are the result of special orders. The products are developed to customer specifications and the price is negotiated. Sometimes flexible credit terms are allowed for the "special order" (typically larger) customers with which Jacobs has a working relationship. The company seems to place a high value on being involved closely with its customers' engineering personnel and appears to trade some credit flexibility for the privilege of being so involved.

The general accounting department at Jacobs consists of several accounting clerks who process receivables and payables. Due to retirements and resignations, the department lost five
employees and has only been able to replace them with two new employees. The department has been understaffed by 30% for most of the year. The five employees who left were highly competent. The new employees are not as well qualified as the previous employees.

The supervisor is responsible for preparing an Accounts Receivable aging schedule and calculating the estimate for bad debts. An account classification indicating account sizes and numbers of customers and an aging schedule are provided on pages 9 and 10 of this packet. The supervisor typically estimates bad debts expense at 1% of annual credit sales and has not deviated from this policy in the current year. Credit sales, on the average, constitute about 80% of total annual sales. Previous years' provisions for bad debts have been reasonably adequate; no audit adjustments were proposed for the allowance account in prior years.

For the 1991 audit, positive confirmations were used. All accounts over $50,000 were confirmed. For 1991, this involved 60 accounts. Ten percent of the remaining accounts were randomly selected to be confirmed. There were a few errors found as a result of last year's confirmation process, but all material errors were corrected to the auditor's satisfaction.

The credit department consists of the credit manager and two clerks who approve credit terms for all customers with sales greater than $500 or sales that would increase a customer's account balance to more than $500. They also review requests for increases in credit limits for previously approved customers. The credit department seems much busier than in previous years. Sales managers seem to be pressuring the department to expedite reviews so they can increase sales to exceed quotas. There are controls in place for credit reviews, but you think that adherence to the controls may have become more lax lately due to increasing pressure from the sales managers. Further, due to the increase in workload from increased sales, the credit department has not been able to maintain the usual level of following up on slow paying customers.

4. Property, Plant and Equipment (PPE)

The cost accounting department, located in the executive offices, does the accounting for PPE. The majority of the PPE account consists of the factories and machines used to manufacture the company's products. Furniture and fixtures are a minor component of the account. The department consists of 3 clerks and a supervisor. Both the clerks and the
supervisor seem to be competent, well trained individuals. The person with the least amount of tenure is the department has been there for four years.

Expenditures for PPE over $30,000 require authorization by the Vice President of Operations. He verifies that the requested expenditure is consistent with the capital budget currently in place. Requests for other expenditures may be authorized by the plant manager. When a requisition is received from the plant manager, a prenumbered purchase order is entered into the computer by one account clerk in cost accounting, checked by another and reviewed by the supervisor. Account classifications are determined by the account clerks and are reviewed by the supervisor. A copy of the purchase order is returned to the plant manager, two copies are filed in cost accounting and three copies go to the vendor. When the goods are received by the plant manager, he sends one copy of the receiving report to the cost accounting department and one copy to the general accounting department. The vendor sends the invoice directly to general accounting. Cost accounting sends the purchase order to general accounting to indicate that the receiving report matches the purchase order and that payment may be processed if the invoice is in order.

The client has prepared a PPE schedule for 1992. This schedule appears on page 11 and a comparable schedule for 1991 appears on page 12 of this packet. Major additions to this account typically occur in the plant machinery and equipment category. Since much of the machinery and equipment is 10 to 12 years old, annual maintenance expenses for these assets have been approximately 5% of the beginning balance (original cost) of the asset category over the past few years. All maintenance expenses for plant machinery and equipment are recorded as manufacturing overhead. Maintenance expenses for 1992 were approximately $1,050,000 (unaudited), approximately $925,000 for 1991 and approximately $802,000 for 1990.
5. Debt Covenants

As a condition of the long-term debt held by the company the following covenants apply:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Ratio:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current Assets</td>
<td>(\geq 1.60)</td>
<td>1.65</td>
<td>1.89</td>
</tr>
<tr>
<td>Current Liabilities:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-term Debt to Equity Ratio:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\frac{\text{Current Maturities + LTD}}{\text{Equity}})</td>
<td>(\leq 1.00)</td>
<td>.35</td>
<td>.41</td>
</tr>
<tr>
<td>Net Worth:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equity</td>
<td>(\geq 12) Million</td>
<td>$21.8) Mill.</td>
<td>$20.2) Mill.</td>
</tr>
</tbody>
</table>

6. Account Classification by Size and Number of Customers

<table>
<thead>
<tr>
<th>Customers' Receivables Balances as of 12-31-92 (unaudited)</th>
<th>Number of Customers</th>
<th>Total Class Balance (000's omitted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$100,000 - $150,000</td>
<td>19</td>
<td>$2,296</td>
</tr>
<tr>
<td>50,000 - 99,999</td>
<td>51</td>
<td>3,583</td>
</tr>
<tr>
<td>25,000 - 49,999</td>
<td>47</td>
<td>1,746</td>
</tr>
<tr>
<td>10,000 - 24,999</td>
<td>68</td>
<td>953</td>
</tr>
<tr>
<td>5,000 - 9,999</td>
<td>204</td>
<td>1,258</td>
</tr>
<tr>
<td>2,500 - 4,999</td>
<td>70</td>
<td>221</td>
</tr>
<tr>
<td>1,000 - 2,499</td>
<td>32</td>
<td>54</td>
</tr>
<tr>
<td>500 - 999</td>
<td>35</td>
<td>25</td>
</tr>
<tr>
<td>0 - 499</td>
<td>31</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>557</td>
<td>$10,143</td>
</tr>
</tbody>
</table>
7. Account Classification by Age of Receivable (000's omitted)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Current</td>
<td>$6,857</td>
<td>67.6%</td>
<td>$5,585</td>
<td>75.0%</td>
</tr>
<tr>
<td>Less than 30 days past due</td>
<td>1,116</td>
<td>11.0%</td>
<td>1,117</td>
<td>15.0%</td>
</tr>
<tr>
<td>30 - 60 days past due</td>
<td>426</td>
<td>4.2%</td>
<td>521</td>
<td>7.0%</td>
</tr>
<tr>
<td>61 - 120 days past due</td>
<td>801</td>
<td>7.9%</td>
<td>149</td>
<td>2.0%</td>
</tr>
<tr>
<td>121 - 180 days past due</td>
<td>517</td>
<td>5.1%</td>
<td>67</td>
<td>0.9%</td>
</tr>
<tr>
<td>More than 180 days past due</td>
<td>426</td>
<td>4.2%</td>
<td>7</td>
<td>0.1%</td>
</tr>
<tr>
<td>Accounts Receivable</td>
<td>$10,143</td>
<td>100.0%</td>
<td>$7,446</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

8. Audit Planning Guidelines

Last year's audit of accounts receivable was budgeted at 40 hours and the audit of property, plant and equipment was budgeted at 10 hours. Planning materiality for the 1992 audit has been set at $325,000.
JACOBS MANUFACTURING COMPANY

Summary of Property, Plant and Equipment and Accumulated Depreciation for the Year ended December 31, 1992

(All dollar amounts are in thousands)

<table>
<thead>
<tr>
<th></th>
<th>Cost</th>
<th>Depreciation</th>
<th>Net Book Value</th>
<th>Depreciation Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1/1/92</td>
<td>Additions</td>
<td>Retirements</td>
<td>12/31/92</td>
</tr>
<tr>
<td>Buildings*</td>
<td>$5,163</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plant Machinery and Equipment*</td>
<td>21,150</td>
<td>$3,571</td>
<td>$212</td>
<td>24,509</td>
</tr>
<tr>
<td>Vehicles</td>
<td>650</td>
<td>102</td>
<td>33</td>
<td>719</td>
</tr>
<tr>
<td>Office Equipment</td>
<td>292</td>
<td>15</td>
<td>10</td>
<td>297</td>
</tr>
<tr>
<td>Furniture and Fixtures*</td>
<td>224</td>
<td>20</td>
<td>26</td>
<td>218</td>
</tr>
<tr>
<td></td>
<td>$27,479</td>
<td>$3,708</td>
<td>$281</td>
<td>$30,906</td>
</tr>
</tbody>
</table>

* Some or all of the depreciation expense for these categories is debited to manufacturing overhead (90% of Buildings, 100% of Plant, Machinery and Equipment, and 70% of Furniture and Fixtures.)
## JACOBS MANUFACTURING COMPANY

Summary of Property, Plant and Equipment and Accumulated Depreciation for the Year ended December 31, 1991

(All dollar amounts are in thousands)

<table>
<thead>
<tr>
<th>Cost</th>
<th>Depreciation</th>
<th>Net Book Value</th>
<th>Depreciation Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1/91</td>
<td>Additions</td>
<td>Retirements</td>
<td>12/31/91</td>
</tr>
<tr>
<td>Buildings*</td>
<td>$5,163</td>
<td>$5,163</td>
<td>$826</td>
</tr>
<tr>
<td>Plant Machinery and Equipment*</td>
<td>18,386</td>
<td>$2,974</td>
<td>$210</td>
</tr>
<tr>
<td>Vehicles</td>
<td>573</td>
<td>148</td>
<td>71</td>
</tr>
<tr>
<td>Office Equipment</td>
<td>285</td>
<td>29</td>
<td>22</td>
</tr>
<tr>
<td>Furniture and Fixtures*</td>
<td>222</td>
<td>26</td>
<td>24</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>$24,629</strong></td>
<td><strong>$3,177</strong></td>
<td><strong>$327</strong></td>
</tr>
</tbody>
</table>

* Some or all of the depreciation expense for these categories is debited to manufacturing overhead (90% of Buildings, 100% of Plant, Machinery and Equipment, and 70% of Furniture and Fixtures.)
Now that you know something about Jacobs Manufacturing Company and its accounts receivable (including the allowance) and property, plant and equipment accounts, think about how you would plan to audit these two accounts for 1992. Your answers to the questions on this page should reflect your beliefs about the relative importance of these two accounts within the particular audit.

First consider the two accounts together:

Which of the two accounts are you more concerned about? (Circle one)

<table>
<thead>
<tr>
<th>ACCOUNTS RECEIVABLE</th>
<th>or</th>
<th>PROPERTY, PLANT AND EQUIPMENT</th>
</tr>
</thead>
</table>

How would you rate each of the accounts on the following scale?

<table>
<thead>
<tr>
<th>10</th>
<th>Very High Concern Account</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Mark an “X” on the scale to indicate how you think the property, plant and equipment account of this client rates on the scale</td>
</tr>
<tr>
<td>8</td>
<td>Mark an “O” on the scale to indicate how you think the accounts receivable account of this client rates on the scale</td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Very Low Concern Account</td>
</tr>
</tbody>
</table>
Now consider only the accounts receivable (and allowance) account. For this page, think about how the level of substantive testing you would plan for Jacobs' accounts receivable and allowance accounts would compare to the level planned for other clients of comparable size. Recall that in the case material, it was stated that your firm had decided not to rely on the internal control system of the client company due to the high cost of testing the system. The reliance or non-reliance on the internal control system can have a significant impact on the level of substantive testing that is planned. It is important that you compare the level of substantive testing you would plan for this account to the level you would plan for other clients' accounts where you also do not rely on the internal control system.

On the following scale, please indicate the level of substantive testing you would plan for Jacobs Manufacturing Company's accounts receivable (and allowance) account. (Please circle a number on the scale.)

-5  -4  -3  -2  -1  0  1  2  3  4  5

<table>
<thead>
<tr>
<th>Minimum level of testing for comparably sized client</th>
<th>Significantly less than average level of testing for comparably sized client</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average level of testing for comparably sized client</td>
<td>Significantly more than average level of testing for comparably sized client</td>
</tr>
<tr>
<td>Maximum level of testing for comparably sized client</td>
<td></td>
</tr>
</tbody>
</table>
Below is a list of audit procedures related to accounts receivable. Please indicate by placing a check mark on the line to the left of each procedure, the ones that you would plan for Jacobs Manufacturing Company for the 1992 audit. At the end of this list, you may add any additional procedures you think would be appropriate.

After you have considered the nature and extent of the testing you would plan, estimate and indicate below the total planned hours for auditing accounts receivable (and the allowance).

Approximate total hours

____ Foot accounts receivable listing and agree to General Ledger and lead sheet.
Obtain client's reconciliation to subsidiary ledger.

____ Review accounts receivable listing for large accounts, credit balances and anything else unusual.

____ Analytical review procedures (Check those you would perform)

___ Review ratio of current period's accounts receivable to sales, compare to last year and consider changed conditions in the current period.

___ Scan trial balance of accounts not selected for confirmation to determine reasonableness.

___ Look for trends in sales by product line and ratios of sales to other meaningful measures (e.g. sales per employee).

___ Perform analytical review on aged listing of accounts.

____ Confirmation of accounts receivable (including follow-up procedures for non responses)

Indicate type of confirmation planned: ______________________________

All accounts over: __________________________

Sample of other accounts - approximate number: ______________________

____ Test cut-off of accounts receivable at year-end.

____ Review documents related to any large unusual sales at year-end for propriety of sale.

____ Discuss collectibility concerns with credit manager.

____ Verify accuracy of aging schedule by tracing a sample of accounts to detail.

____ Review payments received subsequent to the balance sheet date.
Review evidence concerning credit worthiness of any new large customers.

Evaluate the adequacy of the allowance for uncollectible accounts at the balance sheet date.

Analyze activity in the allowance and bad debt expense accounts for the year.

For a sample of sales transactions, examine supporting documentation to ensure validity of the sale.

Test the posting of individual sales invoices to the sales register and the customer's account.

For payments received subsequent to the balance sheet date, trace subsequent collections to customer remittance information for a sample of collections.
Now consider only the property, plant and equipment account. For this page, think about how the level of substantive testing you would plan for Jacobs' property, plant and equipment account would compare to the level planned for other clients of comparable size. Recall that in the case material, it was stated that your firm had decided not to rely on the internal control system of the client company due to the high cost of testing the system. The reliance or non-reliance on the internal control system can have a significant impact on the level of substantive testing that is planned. It is important that you compare the level of substantive testing you would plan for this account to the level you would plan for other clients' accounts where you also do not rely on the internal control system.

On the following scale, please indicate the level of substantive testing you would plan for Jacobs Manufacturing Company's property, plant and equipment account. (Please circle a number on the scale.)

-5----4-----3-----2-----1-----0-----1-----2-----3-----4-----5

<table>
<thead>
<tr>
<th>Minimum level of testing for comparably sized client</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significantly less than average level of testing for comparably sized client</td>
</tr>
<tr>
<td>Average level of testing for comparably sized client</td>
</tr>
<tr>
<td>Significantly more than average level of testing for comparably sized client</td>
</tr>
<tr>
<td>Maximum level of testing for comparably sized client</td>
</tr>
</tbody>
</table>
Below is a list of audit procedures related to property, plant and equipment. Please indicate by placing a check mark on the line to the left of each procedure, the ones that you would plan for Jacobs Manufacturing Company for the 1992 audit. At the end of this list, you may add any additional procedures you think would be appropriate.

After you have considered the nature and extent of the testing you would plan, estimate and indicate below the total planned hours for auditing property, plant and equipment.

**Approximate total hours**

___ Test mathematical accuracy of property, plant and equipment schedule showing beginning balance, additions, retirements and ending balance. Reconcile to lead schedule and to General Ledger.

___ Verify major additions with purchase documents or construction cost records.

___ Compare actual costs of additions with the authorized or estimated amounts; investigate the reasons for any differences.

___ Review major retirements.

___ Review repairs and maintenance expense for reasonableness. Investigate any unexpected changes in repair and maintenance expense.

___ Review status of legal title to assets.

___ Compare current period's depreciation expense to prior year for reasonableness.

___ Inquire:

___ Whether depreciation lives reflect current estimates of useful lives.

___ As to idle or excess assets for which permanent impairment in value may have occurred.

___ Obtain the client's depreciation schedule. Test foot and agree to expense. Recalculate depreciation expense on a sample of assets and project results to compare to expense.
QUESTIONNAIRE

Please complete the following questionnaire after you have completed the Jacobs case.

1. How much auditing experience do you have?
   
   ____________ years and __________ months

2. Please indicate your current rank or classification

   _____ Staff    _____ Senior    _____ Supervisor    _____ Manager
   _____ Senior Manager    _____ Partner

3. Please indicate your experience level with regard to planning audits of the accounts receivable account. Have you planned audit testing for the accounts receivable account on:

   _____ 0 audits    _____ 1 - 3 audits    _____ 4 - 7 audits    _____ 8 - 12 audits
   _____ 13 - 18 audits    _____ more than 18 audits

4. Please indicate your experience level with regard to planning audits of the property, plant and equipment account. Have you planned audit testing for the property, plant and equipment account on:

   _____ 0 audits    _____ 1 - 3 audits    _____ 4 - 7 audits    _____ 8 - 12 audits
   _____ 13 - 18 audits    _____ more than 18 audits

5. How likely do you think it is that there is a material error in the accounts receivable account of Jacobs Manufacturing Company? (Please circle one of the numbers on the line below.)

   ---------------------------
   0 1 2 3 4 5 6 7 8 9 10
   no not Equally very Certain
   chance very likely that likely as not
6. How likely do you think it is that there is a material error in the **property, plant and equipment** account of Jacobs Manufacturing Company? (Please circle one of the numbers on the line below.)

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>no</td>
<td>not</td>
<td>equally</td>
<td>likely</td>
<td>very</td>
<td>certain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chance</td>
<td>very</td>
<td>likely that</td>
<td>there is an error</td>
<td>as not</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The term "consequences" in the next four questions refers to some loss that you or your firm could experience as a result of not detecting a material error, if an error is discovered after the audit report is issued. For example, you could receive a negative performance evaluation, lose your job and/or be judged personally liable in a lawsuit brought by injured parties. Similarly, your firm could experience, for example, the loss of the client, bad publicity or a large legal judgment against it. The questions below ask for your perception of the severity of the negative outcomes that could result for the particular situation in this case, as it relates to each of the accounts.

7. Suppose there were a material error in the **accounts receivable** account of Jacobs Manufacturing Company and it was not detected during the audit. How severe do you think the consequences might be for **you personally**, if the error were discovered after the audit report was issued? (Please circle one of the numbers on the line below.)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>not</td>
<td>very severe</td>
<td>moderately severe</td>
<td>severe</td>
<td>very severe at all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. Suppose there were a material error in the **accounts receivable** account of Jacobs Manufacturing Company and it was not detected during the audit. How severe do you think the consequences might be for **your firm** if the error were discovered after the audit report was issued? (Please circle one of the numbers on the line below.)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>not</td>
<td>very severe</td>
<td>moderately severe</td>
<td>severe</td>
<td>very severe at all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
9. Suppose there were a material error in the \textit{property, plant and equipment} account of Jacobs Manufacturing Company and it was not detected during the audit. How severe do you think the consequences might be for \textit{you personally}, if the error were discovered after the audit report was issued? (Please circle one of the numbers on the line below.)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>not very severe</td>
<td>moderately severe</td>
<td>very severe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>very severe at all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

10. Suppose there were a material error in the \textit{property, plant and equipment} account of Jacobs Manufacturing Company and it was not detected during the audit. How severe do you think the consequences might be for \textit{your firm} if the error were discovered after the audit report was issued? (Please circle one of the numbers on the line below.)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>not very severe</td>
<td>moderately severe</td>
<td>very severe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>very severe at all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

11. What is your past experience with auditor liability issues? (Check all that apply and specify any other relevant experience)

- [ ] Have been involved in _____ lawsuit(s)
- [ ] Have been involved in _____ out-of-court settlement(s) related to auditor liability
- [ ] Have not been involved in either of the above but awareness of the potential for legal losses influences my planning decisions
- [ ] Other (Please specify):

Thank you again for participating in this research. If you would like to receive a summary report of the results of this project, please write your name and address on the enclosed form and hand it in to the contact person separately from your envelope. \textit{Please do not enclose your name and/or address with these materials or with the materials for the other exercise.} It is important that your responses remain anonymous. Thank you!
General Comments

Thank you for agreeing to participate in this research project. Your time and commitment are greatly appreciated. The purpose of this research is to investigate what factors influence the level of substantive testing done on each account in an audit. To achieve this investigation, you will be asked to complete two tasks and a questionnaire. Your responses will be kept confidential and anonymous. So that a valid contribution to auditing research may be made, please answer the questions honestly and in the context of how you would approach an actual audit situation. There are no right or wrong answers to the requirements of these tasks. The interest is only in what influences your professional judgment; what is important to you. Because this research deals with individual judgments, please do not consult with other auditors for any part of these tasks.

After reading this page, please open up the envelope that says "1" on the front. Please read the instructions and do the task that is in the envelope. After you complete the first task, please put all the material for it back in the envelope marked "1" and close the envelope. Next, please open the envelope marked "2" and follow the instructions for the second task. Please do not refer back to the first task while you are doing the second one. After you complete the second task, please put all the material for it back in the envelope marked "2" and place both envelopes together back in the large envelope.
It would be best if you could complete both of the tasks within the same sitting. If that is not possible, you may complete one of the tasks at one sitting and then complete the other task at a second sitting. If you do complete the two tasks at different times, please note on the front of each envelope the date and approximate time that you completed each task.

Thank you for your cooperation and your time.
APPENDIX E
DESCRIPTION OF THE AXIOMS OF CJM

The purpose of this appendix is to describe the major axioms of CJM (simple independence, double cancellation and joint independence) and provide examples using the three factors; probability (P), effort (E), and consequences (C). The two other axioms, distributive cancellation and dual-distributive cancellation are mentioned.

The first axiom that is tested for simple polynomials is simple independence. Satisfaction of this axiom implies that the preference order for the levels of one factor does not depend on the levels of the other two factors.

Example 43:

P is independent of E and C if:

(P2, E1, C1) \geq (P1, E1, C1) \text{ if and only if }

(P2, E2, C2) \geq (P1, E2, C2)

Simple independence for P will be tested for each level of P, and every combination of E and C. Simple independence will hold if the subject always prefers P3 to P2, P2 to P1 and P3 to P1, regardless of the accompanying levels of E and C. 44

---

43The numbers 1, 2 and 3 will follow each factor initial to indicate level. For example, P1 indicates the first level of the probability factor, the description implying "low probability of error." The symbol \( \geq \) should be interpreted as "is preferred or equivalent to."

44In the context of the audit planning judgments being studied, it is assumed that the levels within an factor will be preferred in descending order of level, with level 3 always being preferred to level 2 which will always be preferred to level 1. It is expected that the simple independence axioms will not be violated. The assumption about level preference makes it possible to define dominance and trade-off violations of the joint independence axiom.
The same tests will be done for E and C. Simple independence will hold for effort if the subject always prefers lower effort to higher effort, no matter what the levels of P and C are. Simple independence will hold for consequences if the subject always prefers to plan a higher level of substantive testing for an account that could result in more severe consequences over an account with lesser related consequences if a material error remains undetected, no matter what levels of P and E are included.

Double cancellation is required along with simple independence for any of the four simple polynomials to be valid. This axiom relates to only two variables; the third variable is kept at a constant level for all comparisons between a pair of variables. Double cancellation is similar to the transitivity axiom required for utility theory.

Example:

P and E satisfy double cancellation if:

\[(P_3, E_2, C_3) \geq (P_2, E_1, C_3) \text{ and } (P_2, E_3, C_3) \geq (P_1, E_2, C_3) \implies (P_3, E_3, C_3) \geq (P_1, E_1, C_3)\]

For double cancellation to be testable, the two antecedent conditions must be satisfied first. Nygren (1985) shows through the use of completely random data and simple independence conditioned data that satisfaction of antecedent conditions can provide diagnostic information. The more ordered the data (greater number of factors satisfying simple independence), the greater the number of antecedent conditions satisfied and the greater the number of double cancellation tests that are possible.
Tests of joint independence, including tests for dominance and trade-off violations provide much of the diagnostic information to help determine an appropriate model for the auditor's preferences for substantive testing levels between accounts. If two factors are jointly independent of the third factor, then one combination of the two factors would always be preferred to another combination no matter what the level of the third factor is.

Example:

P and E are jointly independent of C whenever:

\[(P_3, E_3, C_3) \geq (P_2, E_2, C_3) \text{ if and only if } (P_3, E_3, C_2) \geq (P_2, E_2, C_2)\]

In the context of this example, probability and effort would be jointly independent of consequences if an account with high probability of error and low effort would be preferred to an account with medium probability of error requiring a medium effort to increase substantive testing no matter what the level of associated consequences.

A dominant violation of the joint independence axiom would occur if the subject violated joint independence in a situation like the example above. For both P and E, level 3 strictly dominates level 2, so P3, E3 should always be preferred to P2, E2, no matter what the level of C is.

Alternatively:

If \((P_3, E_2, C_3) \geq (P_2, E_3, C_3)\) but \((P_3, E_2, C_1) < (P_2, E_3, C_1)\), this would indicate a trade-off violation of joint independence.

In the context of this trade-off situation, the subject would prefer to plan a high level of substantive testing for an account with a high probability of error and medium effort to increase testing if the severity of consequences is high. But, the subject would prefer to plan high substantive testing for an account with medium probability of error and low effort if there is only a low level of consequences.
The distributive cancellation axiom ideally requires at least a $4 \times 4 \times 2$ design\textsuperscript{45} (Krantz and Tversky 1971), where the outside factor in the distributive model has two levels. For example, if the model were $P(E + C)$, probability would require two levels and effort and consequence would require four levels each.

The dual-distributive cancellation axiom requires at least a $5 \times 5 \times 5$ design, where all factors have five levels. The testing of both of these axioms are complex and the results are not highly diagnostic. For a $3 \times 3 \times 3$ design, such as that used here, the application of CMSCAL and analysis of joint independence violations are expected to provide adequate determination of whether an additive or non-additive model would be more appropriate to represent an auditor's preferences based on probability, effort and consequence. Section 5.2.4.2 describes CMSCAL and the implications of violations of the joint independence axioms.

Figure 8 on the next page presents a flowchart for the diagnosis of simple polynomials. The flowchart relates the axioms described in this appendix to the simple polynomials for which they test (described in Appendix F).

\textsuperscript{45} When a $3 \times 3 \times 3$ design is used, both the distributive cancellation and dual-distributive cancellation axioms are tested in a limited way, since not all possible tests can be performed.
Figure 8. Flowchart for the Diagnosis of Simple Polynomials

(Adapted from Krantz and Tversky (1971))
There are four simple polynomials in three variables that are the alternative models considered in CJM analysis. They are the additive, multiplicative, distributive and dual-distributive models.

The additive and multiplicative models can not be distinguished from one another by axiomatic CJM techniques if scale values are positive. The log of the multiplicative model can be taken to achieve an additive model. Since logarithmic transformation is monotonic, the value of the response variable for the transformed model will maintain the same rank order as the response variable of the original model. Since the factors being considered in this research will have positive scale values, I will discuss only the additive model.

**Additive model.** Example:\(^{46}\):

\[
I_{ijk} = f_1(P_i) + f_2(E_j) + f_3(C_k)
\]

\(I_{ijk} = \) Likelihood of preferring high substantive testing when probability of error is at level i, incremental effort required for high level of substantive testing is at level j, and consequence is at level k.

\(f_1(P_i) = \) scale value for level i of probability factor

\(f_2(E_j) = \) scale value for level j of effort factor

\(f_3(C_k) = \) scale value for level k of consequence factor

---

\(^{46}\)The notation is the same as that used by Schneider (1982, 1984).
An additive model is a compensatory model that has been found in many judgment studies to represent subjects' decisions. The compensatory nature of the model implies that an unfavorable level on one factor can be "compensated for" by a favorable level on another factor. In the context of the factors currently under study, an additive model would imply that if the probability of error is low, a low effort required to increase substantive testing and/or a high level of consequence could motivate the auditor to prefer to increase substantive testing for that account. Similarly, a high effort level required to increase substantive testing could be offset by a high probability of error and/or a high level of consequence if a material error is not detected.

**Distributive Model.** A distributive model implies that one of the factors is more critical to the subject's preference than the other two. The other two factors may be interpreted as a combined or common factor.

Examples:

(a) \[ I_{ijk} = f_1(P_i) [ f_2(E_j) + f_3(C_k) ] \]

(b) \[ I_{ijk} = f_2(E_j) [ f_1(P_i) + f_3(C_k) ] \]

(c) \[ I_{ijk} = f_3(C_k) [ f_1(P_i) + f_2(E_j) ] \]

In example (a), probability is the outside factor. This model implies that an increase in probability of error would increase the preference for increasing substantive testing more than increases in the levels of the other two factors. This type of model would support the auditing literature that emphasizes the importance of the relative frequency of errors. One could also think of effort and consequence as two types of costs to be traded off in considering levels of testing across accounts.
For example (b), effort required for increasing substantive testing is the outside factor. This model implies that a decrease in effort would increase the preference for high audit emphasis more than an increase in probability and/or consequence. This model would reflect the auditor's emphasis on budgetary considerations.

For example (c), consequence is the outside factor. This model implies that an increase in the expected consequence of subsequent detection of a material error would increase the preference for high substantive testing more than an increase in probability and/or a decrease in effort. This model would reflect the auditor's sensitivity toward the potential for legal losses and/or loss of client.

**Dual-Distributive Model.** A dual-distributive model implies that there is an interaction between two of the three factors and the contribution of the third factor is independent of the contributions of the other two. Depending on the range of scale values resulting from numerical CJM, the third factor may be interpreted as being the least important in establishing relative preference for planning a high level of substantive testing for accounts.

Examples:

(a) $I_{ijk} = [f_1(P_i) f_2(E_j)] + f_3(C_k)$

(b) $I_{ijk} = [f_1(P_i) f_3(C_k)] + f_2(E_j)$

(c) $I_{ijk} = [f_2(E_j) f_3(C_k)] + f_1(P_i)$

The model in example (a) implies that high levels of probability and effort would most influence a preference regardless of the level of consequences. Low levels of probability and effort could be offset somewhat by a high level of consequence.

The model in example (b) implies that high levels of probability and consequence would most influence a preference regardless of the level of effort. This model may be reasonable if effort is considered only when the combination of probability and
consequence is at a medium level. In this case, effort may be used as a deciding factor. At very high and very low levels of combined, P and C, the effort factor may be inconsequential.

The model in example (c) implies that high levels of effort and consequence would most influence a preference regardless of the level of probability. Low levels of effort and consequence could be offset somewhat by a high level of probability.
REFERENCES


