INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the original text directly from the copy submitted. Thus, some dissertation copies are in typewriter face, while others may be from a computer printer.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyrighted material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is available as one exposure on a standard 35 mm slide or as a 17" × 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. 35 mm slides or 6" × 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
AUDITORS' PLANNING STAGE MATERIALITY JUDGMENTS AND THE MEDIATING EFFECTS OF LEVEL OF RESPONSIBILITY, FIRM AFFILIATION, AND AUDIT TECHNOLOGY: AN EXPERIMENT

Tidrick, Donald Eugene, Ph.D.

The Ohio State University, 1987

© 1987

Tidrick, Donald Eugene

All rights reserved

U.M.I
300 N. Zeeb Rd.
Ann Arbor, MI 48106
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark √.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print ______
3. Photographs with dark background _____
4. Illustrations are poor copy ______
5. Pages with black marks, not original copy ______
6. Print shows through as there is text on both sides of page ______
7. Indistinct, broken or small print on several pages ✓
8. Print exceeds margin requirements ______
9. Tightly bound copy with print lost in spine ______
10. Computer printout pages with indistinct print ______
11. Page(s) ___________ lacking when material received, and not available from school or author.
12. Page(s) ___________ seem to be missing in numbering only as text follows.
13. Two pages numbered ______. Text follows.
14. Curling and wrinkled pages ______
15. Dissertation contains pages with print at a slant, filmed as received ✓
16. Other ____________________________

________________________________________

________________________________________

U-MI
AUDITORS' PLANNING STAGE MATERIALITY JUDGMENTS
AND THE MEDIATING EFFECTS OF LEVEL OF RESPONSIBILITY,
FIRM AFFILIATION, AND AUDIT TECHNOLOGY: AN EXPERIMENT

DISSERTATION

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

By
Donald Eugene Tidrick, B.S., M.B.A.

* * * * *

The Ohio State University
1987

Reading Committee:                           Approved by:
Thomas J. Burns                               Thomas J. Burns
Bruce R. Gaumnitz                             Adviser
Richard J. Murdock                            Department of Accounting
Jane F. Mutchler
Dennis K. Pearl
DEDICATION

to the memory of my beloved mother,
Hazel G. Tidrick, who died May 30, 1986,
and
to the memory of my friend and classmate,
Daniel R. Lenox, Ph.D. Candidate, who died
July 18, 1986
VITA

May 7, 1953 ................ Born -- Coshocton, Ohio

1976. ....................... B.S., The Ohio State University, Columbus, Ohio

1977. ....................... MBA, University of Wisconsin, Madison, Wisconsin

1978-1980 ................... Staff and Senior Auditor, Peat, Marwick, Mitchell & Co., Columbus, Ohio

1981. ....................... Instructor, Department of Accounting, The Ohio State University, Columbus, Ohio

1982-1986 ................... Teaching Associate, Department of Accounting, The Ohio State University, Columbus, Ohio

1987. ....................... Assistant Professor, Department of Accountancy, University of Notre Dame, South Bend, Indiana

FIELDS OF STUDY

Major Field: Accounting

Minor Fields: Decision Theory (Engineering) Statistics
IV. RESULTS OF ANALYSES.

Introduction
Experience Effects
Firm Effects
Audit Technology Effects
Risk Attitude
Additional Analyses
Comments on Statistical Assumptions
Chapter IV Footnotes

V. REVIEW AND SUMMARY OF RESULTS, CONCLUSIONS, LIMITATIONS, AND EXTENSIONS

Review and Summary of Results
Conclusions
Limitations of the Experiment
Extensions

APPENDICES

A. Instructions and Background Information
B. Attitude Measurement Questionnaire
C. Survey Questionnaire
D. Results of Analyses (Balance Sheet Judgments)

BIBLIOGRAPHY
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>All Significant Cues (p &lt; .05) By Level of Responsibility, By Firm Affiliation</td>
<td>208</td>
</tr>
<tr>
<td>19</td>
<td>Only One Significant Cue (p &lt; .05) By Level of Responsibility, By Firm Affiliation</td>
<td>210</td>
</tr>
<tr>
<td>20</td>
<td>Standardized Balance Sheet Materiality Scores By Level of Responsibility and By Firm Affiliation</td>
<td>288</td>
</tr>
<tr>
<td>21</td>
<td>Summary of Level of Responsibility Effects For the Balance Sheet Judgments</td>
<td>291</td>
</tr>
<tr>
<td>22</td>
<td>ANOVA Table For Firm Effects For the Balance Sheet Judgments</td>
<td>294</td>
</tr>
<tr>
<td>23</td>
<td>Average Variance By Firm and Level of Responsibility Based on Standardized (Balance Sheet) Materiality Scores</td>
<td>298</td>
</tr>
<tr>
<td>24</td>
<td>Summary of Firm Effects For the Balance Sheet Judgments</td>
<td>300</td>
</tr>
<tr>
<td>25</td>
<td>ANOVA Table For Audit Technology Effects For the Balance Sheet Judgments</td>
<td>303</td>
</tr>
<tr>
<td>26</td>
<td>Summary of Audit Technology Effects For the Balance Sheet Judgments</td>
<td>306</td>
</tr>
<tr>
<td>27</td>
<td>All Significant Cues (p &lt; .05) For the Balance Sheet Judgments; By Level of Responsibility, By Firm Affiliation</td>
<td>310</td>
</tr>
<tr>
<td>28</td>
<td>Only One Significant Cue (p &lt; .05) For the Balance Sheet Judgments; By Level of Responsibility, By Firm Affiliation</td>
<td>311</td>
</tr>
</tbody>
</table>
CHAPTER I -- INTRODUCTION

The purpose of this chapter is to introduce the research project along with a brief description of its motivation. In addition, this chapter identifies the organization of the remainder of the dissertation.

The Concept of Materiality

Reference to the professional accounting and auditing literature suggests that, definitionally, there is very little controversy about the meaning of "materiality" as a concept. For example, in its Statement of Financial Accounting Concepts (SFAC) No. 2, the Financial Accounting Standards Board (FASB) commented (FASB [1980, para. 132]):

The essence of the materiality concept is clear. The omission or misstatement of an item in a financial report is material if, in the light of surrounding circumstances, the magnitude of the item is such that it is probable that the judgment of a reasonable person relying upon the report would have been changed or influenced by the inclusion or correction of the item.

The Auditing Standards Board (ASB) of the American Institute of Certified Public Accountants (AICPA) described materiality similarly in its Statement of Auditing Standards (SAS) No. 47 (AICPA [1983, para. 3]):

The concept of materiality recognizes that some matters ... are important for fair presentation of financial statements in conformity with generally accepted accounting principles, while other matters are not important.
efficiency may be to increase the importance which CPA firms place on planning stage activities. Yet, prior research has scarcely addressed the auditor's materiality judgment within a planning context.

This research focuses on auditors' materiality judgments in a planning context. The planning stage materiality judgment (termed "the preliminary judgment about materiality" by SAS No. 47) was chosen for three reasons. First, prior studies have almost exclusively examined auditors' "final" (i.e., evaluation stage) materiality judgments. Yet the professional auditing standards (SAS Nos. 22 and 47) indicate that materiality judgments are made at both the planning and evaluation stages. Hence, it is a legitimate topic for investigation, and the materiality judgment at the planning stage chronologically precedes the judgment at the evaluation stage. Second, insights gained with respect to auditors' planning stage materiality judgments may extend to the evaluation stage. Third, there should be actual auditing ramifications to these judgments at the planning stage. Preliminary judgments about materiality which are, in some sense, "too high" may lead to the gathering of insufficient audit evidence and audits which are not in accordance with GAAS. Preliminary judgments about materiality which are, in some sense, "too low" may lead to audits which are inefficient (i.e., "over-auditing").

In view of prior research involving audit judgment, it seems likely that the participants will differ in their specified planning
A secondary issue explores the extent to which differences among auditors in their planning stage materiality judgments may be attributable to differing "risk attitudes." This concept of risk attitude deals with the tradeoff of Type I and Type II errors (i.e., issuing a qualified audit opinion with respect to a client's financial statements that contain no "material" misstatements and issuing an unqualified audit opinion with respect to a client's financial statements that do contain "material" misstatements, respectively) that may be viewed as implicit in rendering materiality judgments at the aggregate financial statement level. The purpose is to identify factors associated with differing materiality judgments in order to explore reasons why the participants may differ in their stated materiality thresholds. If auditors differ systematically in their judgments due to different perceptions of the various consequences associated with these so-called alpha and beta risks, then evidence may be obtained that different materiality thresholds are consistent with the notion of auditors' utility maximization.

Objectives of the Dissertation

The three principal objectives of this dissertation involve the identification of factors hypothesized to be associated with systematic differences in consensus among auditors' planning stage materiality judgments. (As mentioned, consensus is examined in terms
particular, the lens model applies to human judgments which are based on a set of cues related to the environment; these cues are related in probabilistic terms to some criterion variable. Because no criterion exists with which to assess the accuracy of auditors' materiality judgments (for either the planning or evaluation stage), consensus was emphasized.\textsuperscript{6}

Six of the Big Eight firms participated in this study. Three firms were viewed as highly structured (and were classified as "structured" in this study) and three were viewed as less structured (and were classified as "unstructured" in this study) based on Kinney [1986]. Each firm provided twelve auditors as participants--four each at the senior, manager, and partner levels. Thus, a total of 72 subjects participated. Each subject was asked to specify the appropriate planning stage materiality threshold at the aggregate financial statement level, separately for the income statement and balance sheet, for each of the 24 unique combinations of income (loss) before taxes, net sales, and total assets. Four cases were repeated in order to measure test-retest consistency.

Contributions

A major contribution of this research concerns the measurement of consensus. This research focuses on two aspects of consensus in order to assess the agreement among auditors in their materiality judgments: (1) tests of the significance of the differences in location
the evaluation stage; and, as a result, it may be a more interesting task with which to explore auditors' judgments.

In addition, prior research has reported unclear results about the effect of experience on the consensus among auditors in their judgments, in general, and in their materiality judgments, in particular. This research examines the impact of level of responsibility (presumably related to experience) on the consensus among practicing auditors in the context of their planning stage materiality judgments. Further, insights may be gained about the effect of level of responsibility on these materiality judgments under different audit technologies.

This study will identify whether there are firm differences among the participating Big Eight CPA firms associated with their planning stage materiality judgments. In addition, any effects attributable to different audit technologies (regarding highly structured versus less structured firms) on the consensus of auditors' preliminary judgments about materiality will be identified. Although prior research has reported somewhat mixed results about Big Eight versus non-Big Eight or national versus non-national firm differences, relatively little is known about whether differences exist across Big Eight firms. Similarly, little is known about the effect which different audit technologies may have on auditors' judgments.

Another potential contribution of this research relates to the acknowledgment that auditors may rationally differ in their
example, if materiality thresholds are significantly affected by risk attitudes, then any guidelines would have to provide enough latitude to allow for such differences. Otherwise, one could expect strong opposition to any proposed guidelines.

Organization of the Dissertation

The remainder of the dissertation is organized in four chapters. Chapter II reviews the applicable literature. The major relevant pronouncements of the standard-setting bodies are presented first. The empirical studies of auditors' materiality judgments are discussed second. Other relevant research is reviewed third, including the applicable literature from psychology relative to the lens model and the role of attitudes. Chapter III identifies the particular research interests and hypotheses to be investigated. Further, Chapter III discusses the planned research approach, including the general methodology as well as the specific analyses to achieve the stated research objectives. Chapter IV reports the results of the statistical analyses. Finally, Chapter V offers a brief summary and some conclusions, identifies the limitations of the research, and presents suggestions for future research.
3. Audit technology will be discussed in more depth in Chapters II and III. This dissertation investigates the differences between the highly structured firms and the less structured firms, based on Kinney [1986]. Kinney classified the Big Eight firms as "structured," "intermediate," or "unstructured." This dissertation dichotomizes the six participating Big Eight firms along these lines, except that one firm classified as "intermediate" by Kinney was treated as "unstructured." Review of the participating firms' audit manuals suggested that, with respect to these planning stage materiality judgments, this particular firm was similar to the other firms classified as unstructured.

4. The role of attitudes is discussed in detail in Chapter II. The methodology employed with respect to attitudes has support in the psychology and accounting literatures. Kida [1980] provided the basis for this extension of attitudinal differences into the domain of auditors' materiality judgments. Kida focused on auditors' going concern qualifications and showed that different attitudes with respect to Type I and Type II errors could affect participants' decisions to issue such qualifications in an experimental context.

5. As stated in Chapter II, the descriptions of the lens model in this dissertation rely on the discussions provided by Ashton [1982b] and Dudycha and Naylor [1966].

6. As discussed in Chapter III, Ashton [1982a, p.75] indicates that consensus, self-insight, and judgment stability are often used when a criterion (such as accuracy or normativeness) is unavailable. Prior research efforts, discussed in Chapter II (e.g., Messier [1983] and Krogstad, et al. [1984]), have reported relatively low consensus, but relatively high self-insight and stability on materiality tasks. Since this study emphasizes comparisons among groups of auditors, rather than within individual auditors, consensus is of primary interest.
The general nature of the materiality judgment and the related issue of specific criteria or guidelines were among the first issues placed on the FASB's agenda. The FASB [1975, paragraph 41] stated, "In placing the topic of materiality on its agenda, the Board tentatively accepted the view that it would be desirable to establish criteria for the determination of materiality in financial accounting and reporting."

The outcome of the discussion memorandum was ten paragraphs pertaining to materiality in the text of Statement of Financial Concepts (SFAC) No. 2 (FASB [1980]). However, SFAC No. 2 contained no quantitative guidelines. Despite their observation (paragraph 163) that "The courts have stepped in to fill the gap," the Board concluded (paragraph 131) that their "... present position is that no general standards of materiality could be formulated to take into account all the considerations that enter into an experienced human judgment."

The Board conceded that it may reconsider the appropriateness of issuing general quantitative materiality guidelines or, alternatively, guidelines applicable to specific standards on a piecemeal basis.¹

In 1978, the ASB issued Statement on Auditing Standards (SAS) No. 22, entitled "Planning and Supervision." SAS No. 22 identified several items to be considered when planning the audit (AICPA [1978, paragraph 3]). Of particular interest here was their reference to "preliminary estimates of materiality levels for audit purposes."
"Although the auditor should be alert for errors that could be qualitatively material, it ordinarily is not practical to design procedures to detect them." The planning stage materiality judgment hinges primarily on quantitative factors, whereas the evaluation stage materiality judgment depends to a greater extent on qualitative factors. Even though the two materiality judgments theoretically would be the same if the information at the evaluation stage were available for planning, as a practical matter, the judgments will differ.

In summary, the SAS requires auditors to explicitly consider materiality in audit planning. Although materiality is not required to be quantified, Zuber, et al. [1983, p. 42] point out that "... quantification is the most practical way to consider such an estimate in audit planning." SAS No. 47 provides a broad conceptual discussion of the impact of materiality on the conduct of an audit, and clarifies the earlier reference to "preliminary estimates of materiality" in SAS No. 22. However, it contains no quantitative criteria or guidelines to assist practitioners in their materiality judgments at either the planning or evaluation stages.

Recently, a practicing CPA expressed his firm's desire for specific materiality criteria in "Letters to the Editor" of the Journal of Accountancy (Godick [1983, p. 101]). He observed that "... the nitty-gritty dollar determination of materiality hasn't yet been addressed." The lack of agreement about such guidelines can be seen
Australian Society of Accountants, jointly issued a Statement of Accounting Standards on "Materiality in Financial Statements." A similar pronouncement was issued by the New Zealand Society of Accountants in 1977. Although these pronouncements differed in their authoritativeness and emphasized the importance of professional judgment in differing degrees, each identified factors to be considered in making materiality judgments, and each presented quantitative rules of thumb.

The Accountants International Study Group [1974] issued a report on the materiality-related practices of Canada, the United Kingdom, and the United States. The Study Group concluded (paragraph 30), "Quantitative guidelines within broad parameters are usually used in practice and we believe they could be developed by the authoritative bodies in the three nations."

David Selley, Auditing Standards Director of the CICA, discussed several of the major issues and historical developments related to materiality judgments (Selley [1984]). With respect to the materiality policies of 16 of the largest Canadian accounting firms, Selley reported (p. 18):

... 11 of the 16 provide quantitative guidelines to their auditing professional staff for determining what is material and that all 11 of them use as the primary or as one among several presumptive guidelines 5% to 10% of normal pre-tax net income for most profit-oriented clients. Two more firms have guidance material in process.

In summarizing the world wide development of materiality criteria, Selley commented (p. 18), "With the exception of the U.S. profession
Empirical Research Regarding Materiality

Holstrum and Messier [1982] classified research regarding materiality as questionnaire survey, archival, or judgment-capture experiments. The following summary of the literature focuses on that research dealing with materiality judgments of auditors (even though these studies have tended to deal with accounting rather than auditing contexts). The earliest research into materiality judgments tended to be the questionnaire survey type (e.g., Woolsey [1954a and b]). In addition, some of the more recent, practice-oriented studies have been of the questionnaire survey type (e.g., Dyer [1975] and Pattillo [1976]). These studies are summarized in Table 1.

These questionnaire survey studies of accountants' and auditors' materiality judgments have generally shown consistent results (although the results of such surveys have not been amenable to statistical analysis). Relative amounts have been more important than absolute amounts, and the most important single factor has been the particular item in question as a percentage of net income (or income before taxes).

Woolsey [1954a] surveyed several hundred individuals and different business organizations as to factors affecting the classification of three particular items (gain from the sale of fixed assets, write-off of unamortized bond discount, and loss due to an earthquake casualty loss) as material versus immaterial. Woolsey reported four conclusions: (1) relative amounts were more important
Table 1 (continued)


Subjects: Approximately 700 respondents from public accounting, industry, and user groups (response rate was about 45%).

Context of Judgments: 1. Extraordinary items; 2. contingencies; 3. accounting changes; 4. policy disclosures; 5. business segments; 6. cumulative judgments; and 7. interim disclosures.

Conclusions: 1. Groups varied, but the auditors tended to be the most conservative; 2. "determinant criteria" (quantitative factors) formed the basis for tentative judgments and "modifying criteria" (qualitative factors) influenced the final judgments; and 3. the item as a percentage of net income was the most influential of 7 determinant criteria considered.
Woolsey's. The subjects were classified as national or non-national. They were asked to identify the single most important factor relative to which the amount in question would be compared for purposes of determining materiality. Then they were asked to classify each situation as material or immaterial for each set of cues. Materiality "thresholds" were inferred based on the most important factor identified and the lowest value of that factor for which the subject decided that the item in question was material.

Dyer concluded that the clearly most important factor in the materiality judgment was the relationship of the item to net income--similar to Woolsey. However, the materiality "thresholds" were substantially lower for all subjects on average and there were no clear-cut differences between national and non-national subjects. Dyer attributed this difference with respect to Woolsey's study to environmental changes (such as the legal environment) over the intervening twenty years.

Pattillo [1976] surveyed nearly 700 subjects representing industry, public accounting, various user groups and academia. A series of 28 cases (each containing six levels of quantitative cues) were mailed to the participants. For each judgment situation (involving unusual and extraordinary items, contingencies, accounting changes, accounting policy disclosures, business segments, cumulative materiality judgments with both similar and dissimilar items, and interim report disclosures) the subjects were asked to identify the
Item as a % of net income 1.75 Moderate to high
Absolute dollar amount 1.92 Moderate to high
Effect on E.P.S. 2.04 Moderate
Item as a % of equity 2.70 Low to moderate
Item as a % of assets 2.88 Low to moderate
Item as a % of liabilities 3.22 Low to none
Item as a % of sales 3.32 Low to none

Much of the materiality research in the 1960s and early 1970s tended to be of the "archival" type (e.g., Bernstein [1967]; Neumann [1968]; Copeland and Fredericks [1968]; and Frishkoff [1970]). In addition, Morris, Nichols, and Pattillo [1984] was a recent archival study of auditors' materiality judgments. These studies are summarized in Table 2.

These studies examined actual financial statements in search of a pattern indicative of the materiality criteria applied to those financial statements. Such studies have been unsuccessful in detecting a "dividing line" (based on the item in question as a percentage of net income) between immaterial and material items. Bernstein and Neumann interpreted their results as evidence of a lack of consensus among auditors in the determination of materiality.

Bernstein [1967] examined a sample of 324 income statements containing extraordinary gains or losses. (The accounting treatment of such items was specified at that time in Accounting Research Bulletin No. 43, Chapter 8, and required material extraordinary items to be excluded from income.) Bernstein attempted to infer materiality in relation to net income based upon any apparent "dividing line" between the magnitude of extraordinary items included in income versus those items which were excluded. In other words, an extraordinary
Table 2 (continued)

Conclusions: 1. Only 3 of the 17 variables considered were significant at the .10 level (the item as a percent of net income, the amount of net worth, and a dummy variable which identified reclassifications from more consequential changes); and 2. the item as a percent of net income was not a good predictor of qualification decisions.

5. Morris, Nichols, and Pattillo [1984]

Issue: Whether consistency qualifications with respect to mandatory changes in interest capitalization policies (SFAS No. 34) were driven by quantitative financial statement relationships.

Sample: 221 annual reports for 1979/80 or for 1980/81 affected by SFAS No. 34.

Conclusion: No pattern emerged and, as a result, consensus among the auditors was inferred to be low.
principles (of a comparable nature) in 1954 and again in 1964. By choosing changes affecting many firms across industry lines, Neumann argued that the "qualitative aspects" of materiality were effectively held constant. Hence, the determination of materiality hinged upon its "quantitative dimensions."

The audit reports were examined for those firms which reported such a change in accounting principle in 1954 and/or 1964. An audit opinion which was qualified as to consistency was presumed to imply a material change in accounting principle, whereas an unqualified opinion was construed to imply immateriality. The magnitude of the effect of the change in accounting principle relative to net income (including the change as reported) was used to determine whether any "dividing line" as to materiality versus immateriality could be observed in the application of the consistency reporting standard. As did Bernstein before him, Neumann concluded that the disclosure decision under consideration did not vary systematically with his measure of materiality, and added (p. 12), "It is my opinion, as the result of this study, that the failure to more specifically define the dimensions of materiality has led to a lack of consensus as to the 'when' and 'how' of the implementation of the consistency standard."

Neumann argued that materiality should not be a criterion for the consistency qualification.

Copeland and Fredericks [1968] developed a measure for the extent of disclosure across sets of financial statements and related that
concluded that there was a positive association between materiality and disclosure, although the coefficient was not significant at the .05 level.

Frishkoff [1970] focused on consistency qualifications in his study of auditors' materiality judgments, similar to Neumann [1968]. However, whereas Neumann considered only one variable, Frishkoff considered seventeen. In addition, Frishkoff introduced multiple discriminant analysis in the study of materiality.

Frishkoff reviewed approximately 2200 annual reports for 1963. His sample was comprised of 190 reports involving financial statements containing changes in accounting principles. Of these, 130 contained qualified opinions and 60 were unqualified. Again, it was presumed that qualification implied materiality.

Using step-wise discriminant analysis, Frishkoff found that only three independent variables were significant at the .10 level. These were: the absolute value of the effect divided by net income, the amount of net worth, and a dummy variable distinguishing reclassifications from other changes. The last two were significant at about the .065 level, whereas the first was significant at the .025 level.

Based upon the single independent variable, percent of net income, Frishkoff attempted to identify a "dividing line" as to immaterial versus material amounts. He inferred that this distinction occurred at about 25 percent of net income. This was surprisingly
and values greater than Y percent are always material. Instead, there appears to be one big area of overlap, regardless of which factor is investigated.

Therefore, the authors concluded that there is an absence of consensus among auditors in these materiality judgments. In addition to suggesting additional research to develop general materiality guidelines, the authors recommended that a company's management disclose its materiality criteria in its financial report, in order to facilitate the comparison of financial statements across companies.

Most of the recent empirical research into auditors' materiality judgments has been of the judgment-capture experiment type (e.g., Boatsman and Robertson [1974]; Moriarity and Barron [1976] and [1979]; Ward [1976]; Newton [1977]; Hofstedt and Hughes [1977]; Firth [1979]; Schultz and Reckers [1981]; Emery, Thakkar, and Moriarity [1982]; Messier [1983]; and Krogstad, Ettenson, and Shanteau [1984]). The participants in the studies were asked to express their materiality judgments based on the case materials presented. These studies are summarized in Table 3.

Boatsman and Robertson [1974] used multiple discriminant analysis to model the materiality judgments of groups of CPAs and securities analysts. Rather than examining actual financial statements in conjunction with the auditors' observed opinions, the authors developed a set of hypothetical cases for which individual judgments were required.
Table 3 (continued)

Experimental task: Identify preferences among different sets of "gambles" (for 3 different contexts); these involved different dollar amounts. Probabilities were elicited with respect to qualification decisions.

Conclusions: 1. Most subjects were consistent as to risk-seeking or avoidance behavior across different contexts (most were risk averse); and 2. subjects tended to use probabilities (at least implicitly) in their decisions to qualify.

5. Hofstedt and Hughes [1977]

Subjects: 19 MBA students.

Experimental task: Whether a loss on the write-down of a subsidiary should be classified as an extraordinary item (on a 100-point scale).

Conclusions: 1. The explanatory ability of the ANOVA models decreased when individuals were aggregated; 2. simple models (reflecting only main effects) had about as much explanatory ability as more complicated models (including interaction terms); and 3. individual differences were conspicuous in several respects.

6. Morarity and Barron [1979]

Subjects: 8 audit partners from 1 national CPA firm.

Experimental task: Specify the planning stage materiality threshold for each of 30 cases.

Conclusions: 1. The task was unfamiliar to the subjects and the experimental materials lacked "mundane reality;" and 2. the poor consensus was influenced by the diversity of the partners' backgrounds.

7. Firth [1979]

Subjects: 150 people--30 each from 3 Big Eight CPA firms, industry, and user groups.

Experimental task: Identify whether an extraordinary item warranted separate disclosure for each of 30 cases.
Table 3 (continued)


Subjects: 10 audit partners, 11 audit seniors (the auditors were from 5 Big Eight firms), and 11 accounting students.

Experimental task: Identify the materiality of a proposed adjusting entry (on a 100-point scale) for each of 32 cases.

Conclusions: 1. Effect on net income was the dominant variable for the auditors; 2. no differences were identified between the seniors and partners; 3. the students' judgments differed from the auditors' judgments; and 4. at least 1 qualitative variable was significant for about 2/3 of the auditors.
employed to generate values for the variables having the mean values observed from the actual reports while also having a variance-covariance matrix consistent with the actual reports. The technique generated data in the form of the multivariate normal distribution.

The subjects were eighteen partners from national CPA firms and fifteen securities analysts. Their task was to sort the 30 cases into three disjoint sets corresponding to the desired reporting treatments (i.e., line-item disclosure in the body of the financial statements, footnote only disclosure, and no disclosure of any kind). In the sense that the subjects were instructed to classify the cases as to how they should be treated, the task was normative.

Based upon the 990 observations (30 cases for each of 33 subjects), three multiple discriminate models were determined, one for each possible disclosure treatment. Models were not prepared for individual decision makers, but rather aggregated all subjects' judgments. The models were then used to predict classifications for each case for each individual. The models accurately predicted 63 percent of the classifications as actually made by the subjects. Further, the authors reported that, if line-item disclosure and footnote disclosure categories were combined as indicating materiality and no disclosure interpreted as immateriality, the model was 84 percent accurate in distinguishing material from immaterial cases.

One variable, the absolute value of the item divided by current net income, contributed 73 percent of the predictive accuracy. The
deemed relevant to the decision; (2) the materiality thresholds; (3) the scale values of the variables affecting the decision; and (4) the basic forms of the decision models." Suggesting that other researchers had considered issues (1) and (2), the authors indicated that their investigation of how individuals make materiality judgments would focus on issues (3) and (4).

They used an approach known as conjoint measurement (CJM) to explore the apparent structural form (i.e., linear, multiplicative, or some combination) of the models used by decision makers in the context of a particular materiality judgment. The CJM technique was different from earlier methodologies employed to investigate materiality judgments in that only ordinal characteristics of materiality decisions were needed for CJM. This had the effect of segregating the issue of materiality thresholds from the analysis. The contributing variables were also treated as ordinal and, hence, there was no assumption as to the linearity of the contributions based on the size of the variables.

The authors selected three variables based upon their review of the applicable literature; they noted that, as a practical matter, CJM was limited to three variables, despite the existence of procedures to mitigate that restriction. The three variables they selected were net income, earnings trend, and firm size. A factorial design was used involving three levels of income, three levels of asset size, and two
appeals to intuition, and (3) this study suggests its predominance in use."

Ward [1976] focused on (p. 138) "... the degree of consensus among professional accountants concerning perceptions about the relative importance of factors associated with issues such as materiality." The subjects were 24 audit partners or managers with national CPA firms. The methodology employed was the Q-sort technique, which, similar to CJM described above, required subjects to specify only rankings. The subjects were provided with information about a hypothetical company's internal controls and a set of financial statements. The subjects were asked to consider the scenario of an error within the valuation of accounts receivable having a corresponding effect on net income (first, overstatement and second, understatement situations were considered).

The author provided 20 index cards expressing various implications to the auditors and to the client firm, given that the error affecting accounts receivable and net income was not detected by the auditors. For example, these included "degree to which audit complied with professional standards," "personal liability," and "loss of future engagements with the client," among others. The task required the subjects to sort the 20 cards into one of five different categories as to the degree of importance of the individual items described on each card. In an attempt to approximate a normal distribution, Ward directed the subjects to sort the cards into stacks
Newton [1977] employed cardinal utility analysis in an exploration of the effect of uncertainty upon CPAs materiality judgments. The purpose was to identify and explain the effect of risk attitudes on materiality decisions. Newton hypothesized (p. 98) that "... the decision an accountant reaches concerning the materiality of an item is related directly to the degree of uncertainty about the issue and the accountant's acceptable [Beta] risk" (i.e., issuing misleading financial statements). As risk aversion increases, materiality thresholds are predicted to become smaller.

Cardinal utility theory was used to assess the risk attitudes of the CPA participants. The subjects were nineteen partners of CPA firms encompassing "Big Eight," regional, and local firms. The subjects were required to express preferences as to sets of "gambles" with specified outcomes and designated probabilities. Based upon a subject's preferences, utilities were inferred.

Three unrelated cases were used. The first was a general case to assess each subject's basic attitudes toward risk outside of an accounting context (called the "standard gamble").

The second case involved marketable securities. Given that a client company owned a portfolio of marketable securities costing $25 million, the auditors were asked to specify the amount of error in the valuation of the securities (relative to the company's $50 million net income) which would cause the auditor to qualify the audit report. Probabilities were introduced by asking the subjects to specify the
materiality guidelines should include explicit consideration of the risk factor.

Hofstedt and Hughes [1977] used a research approach known as the "clinical judgment" paradigm to investigate the relationship of materiality and disclosure decisions. Two criteria were used to evaluate a subject's performance. One, "convergence," concerned whether a given decision maker tended to make consistent judgments for different sets of data. The other, "self-insight," dealt with how well the decision maker was able to describe his or her decision processes and the extent to which relevant cues were used in accordance with the described processes. These two criteria were used to evaluate the information-processing characteristics associated with the judgment task.

The subjects were nineteen second-year MBA students, concentrating in accounting, at Cornell University. The task was to decide whether a loss due to the writedown of a subsidiary should be classified as extraordinary based upon its relationship to the experiment's variables.

The experimental design was a $3^3 \times N$. There were three independent variables--the loss was expressed as a percent of operating income, as a percent of the parent company's total investment in all its subsidiaries, and as a percent of the net book value of the subsidiary under consideration. Each independent variable had three levels representing "low," "medium," and "high." In addition, subjects were
complexity of information processing differed widely. Some subjects had as many as five significant factors (main effects plus interactions) whereas others had as few as one. Further, the cue weights appeared to differ.\footnote{Moriarity and Barron [1979] followed up their earlier article. Their approach was a modified version of their earlier approach; the authors expressed their view (p. 114) that "... the primary contribution of the paper is ... an illustration of a methodology having significant implications for accounting research." However, because the experiment lacked "mundane reality," they conceded that the results of the study might not apply to materiality judgments in practice.}

The approach used, called "conjoint analysis," had the following characteristics: assumed additivity, a fractional factorial design, and part-worth functions. The model assumed to apply was

\[ M = f_1(x_{j1}) + f_2(x_{j2}) + f_3(x_{j3}) + f_4(x_{j4}) + f_5(x_{j5}) \] (2.1)

where:

- \( M \) = the amount judged material
- \( f_i(x_{ji}) \) = the effect of variable \( i \) at level \( j \)

and \( f_1, f_2, ..., f_5 \) were the part-worth functions applicable to the five variables. It was assumed that five factors affecting materiality could be found. The variables chosen were (1) net income, (2) debt-equity ratios, (3) number of shares, (4) total assets, and (5) earnings trend. Total assets was the most important variable for
Firth [1979] compared the disclosure judgments of auditors, managers, and user groups (i.e., securities analysts and bankers) in the United Kingdom. Again, disclosure was used as a surrogate for materiality. The task required subjects to specify whether the extraordinary item warranted separate disclosure for each of 30 cases (half dealt with losses and half dealt with gains of different amounts). A total of 150 subjects participated in the study--30 from each of three Big Eight CPA firms, 30 from industry, and 30 financial statement users. Firth manipulated the gains and losses as a percentage of the following variables: income before extraordinary items, net assets, total assets, market capitalization, sales turnover, and current assets.

The gain or loss distinction was not significant, but the items as a percentage of income before extraordinary items, net assets, total assets, and market capitalization were significant. The variable with the greatest explanatory power was the item as a percentage of income before extraordinary items. The results indicated that the users had lower materiality thresholds than the other groups (which were not significantly different). Even though the judgments of the three CPA firms involved were not significantly different, Firth concluded (p. 291) that "... some consideration should be given to the idea of producing guidelines and yardsticks for materiality-disclosure decisions."
The "choice shift" factor (i.e., the comparison of the pre-group, group, and post-group measurements) was not significant. Although the group process did not appear to significantly shift the mean responses, the group setting did appear to reduce the variation in the responses and provide the subjects with a higher level of confidence in their judgments.

Emery, Thakkar, and Moriarity [1982] used conjoint measurement (CJM) to compare the materiality judgments of auditors and financial statement users. According to the authors (p. 5), their paper was intended to answer three research questions:

First, do the materiality judgments of users of accounting information differ from those of auditors? Second, among users, do the materiality judgments of users with an equity point of view (financial analysts) differ from those of users with a debt point of view (commercial bank loan officers)? Third, within each professional group (i.e., auditors, financial analysts, and loan officers), how consistent are the materiality judgments?

Similar to Moriarity and Barron (M&B) in their 1976 paper, the authors applied axiomatic CJM to identify subjects whose judgments could be modeled as additive. Then numerical CJM was applied to determine the coefficient of the applicable cues in the linear models associated with these subjects. The experiment was virtually a replication of M&B's, except for the use of different subject groups.

Again, eighteen sets of financial statements were prepared representing a full factorial experimental design where the three independent variables (income level, earnings trend, and asset size) had three, two, and three different levels, respectively. The
However they cautioned (p. 7) that their results should be interpreted carefully:

That is, this study provides evidence that auditors and users agree that, in an ordinal sense, a specific item is more material for a firm with $50,000 in income than for a firm with $75,000 in income (all else equal). But the study does not represent evidence about whether the groups agree, in a nominal sense, that the item is or is not material in either case.

They argue that studies which observe differences among various subjects in materiality judgments tend to be studies which focus on the determination of materiality thresholds (e.g., Woolsey [1973b] and Pattillo [1975]).

Messier [1983] examined the joint materiality and disclosure decision for 29 audit partners who were classified as to their levels of experience and their firm type. Experience was differentiated as greater than 20 years of experience, sixteen-20 years, or up to fifteen years. The number of partners in these categories were six, eight, and fifteen, respectively. "Firm type" simply dichotomized "Big Eight" versus "non-Big Eight." The number of partners in these categories were 22 and seven, respectively.

The experiment involved a multivariate lens model framework and required each subject to make two judgments regarding a writedown of inventory in the amount of $1 million: (1) rate the materiality of the writedown on a seven-point scale; and (2) identify the probability that the item should be separately disclosed in the income statement in accordance with APB Opinion No. 30 (on a 100-point scale).
Although it is difficult to make generalizations due to sample size limitations, this latter result is consistent with the hypothesis that, on average, a year of Big Eight experience is more concentrated than a year of non-Big Eight experience. It may be that "experience" should reflect more than simply a chronological dimension.

Using MANOVA, the author reported that both main effects (i.e., experience and firm type) and the experience-firm type interaction were significant at .01. A separate ANOVA was calculated for both the materiality and disclosure decisions. For the materiality judgment, only firm type was significant at .01. For the disclosure judgment, experience, firm type and the experience-firm type interaction were significant.

Concerning consensus, Messier reported that the less-than-fifteen-years-of-experience group had less agreement within the group than did the other two groups. Also, the non-Big Eight group had less agreement than the Big Eight group. As to self-insight, the author concluded that experience appeared to have no effect on self-insight and firm type may have had only minor effect. Stability did not appear to be influenced by either experience or firm type.

Messier summarized that experience and firm type had an effect on the consensus of the subjects' judgments (based on MANOVA), but no effect on self-insight or stability. This was interpreted as being attributable to the complexity of the task. Consequently, Messier suggested (p. 618) that "... researchers interested in studying
In addition, they included three financial (quantitative) variables:

(6) the effect of the proposed adjustment on the current ratio (a reduction of 2.5 to 2.1 vs. a reduction of 2.1 to 1.7), (7) the effect of the proposed adjustment on net income (a decrease of 2.7 percent vs. a decrease of 7.3 percent), and (8) the effect of the proposed adjustment on the earnings-per-share (EPS) trend (no reversal of an upward EPS trend vs. reversal of an otherwise upward EPS trend).

The study considered professional experience by distinguishing the subjects as audit partners ("very experienced"), audit seniors ("moderately experienced"), and accounting students ("inexperienced"). There were eleven partners, ten seniors (the auditors were from five of the Big Eight firms), and eleven students. The authors hoped that including accounting students would be helpful in evaluating the extent to which students may serve as surrogates for professionals in research situations.

The authors prepared 32 hypothetical cases (sixteen separate cases were each repeated). The experimental design was a $3 \times 2^8 \times 1/16$ fractional factorial where the main effects were: (1) experience (three levels—a between-subjects variable) and (2) eight cues (two levels each—all within-subjects variables).

The experimental task required each subject to assume the role of the in-charge accountant on a hypothetical audit engagement. A judgment was needed as to the materiality of a proposed adjusting journal entry affecting the "Allowance for Doubtful Accounts." Subjects were asked to identify the materiality of the adjustment on a
Of the 21 auditors included as subjects, fourteen used one or more of the contextual (nonfinancial) cues. The partners, on average, used 1.6 contextual cues versus 1.7 financial cues. Seniors, on average, used 1.1 contextual cues versus 1.6 financial cues. Of the eleven students in the study, ten used at least one contextual cue and, on average, students used 1.5 contextual cues versus 1.0 financial cue.

The variable "primary users" related to the distinction between public corporations (whose primary users were assumed to be stockholders) versus private companies (whose primary users were assumed to be creditors). It was interesting to note that the variable was significant for only four of the ten partners. Of those partners who used it, six to seven percent of the variance was explained solely by this variable. However, this variable was significant for only two of the eleven seniors, and explained four percent of the variance for one senior and one percent for the other. This variable was significant for only one of the eleven students, but explained 29 percent of the variance for that subject. It is not
Other Relevant Research

The Lens Model

Egon Brunswik developed a framework for representing human judgment known as the "lens" model. The lens model has been summarized by a number of authors, including Ashton [1982b] and Dudycha and Naylor [1966]. Their descriptions are relied on here.

As suggested by Figure 1, the lens model partitions the decision or judgment situation into two parts. The first part (the left-hand side of the model) represents the environment. \( Y_e \) denotes the criterion or "distal" variable and represents the actual outcome. \( X_i \) denotes cue \( i \) representing an item of information which can be used in some respect to predict the criterion variable. Brunswik viewed the cues as being linked probabilistically to the uncertain criterion event. Hence, the decision maker has to base his or her judgment on a subset of the total information set available to the environment. The second part of the model (the right-hand side) represents the individual's judgment system. \( Y_s \) denotes the individual's judgment or prediction. The cue validity coefficients are denoted by \( r_{ei} \) and represent the relationship between \( Y_e \) and each \( X_i \). The cue utilization coefficients are denoted by \( r_{si} \) and represent the extent to which the individual utilizes each \( X_i \) to formulate \( Y_s \).
Using linear multiple regression, the left-hand side of the lens model can be represented by the following:

\[ \hat{Y}_e = b_{e1}X_1 + b_{e2}X_2 + \ldots + b_{en}X_n \]  

(2.2)

where \( \hat{Y}_e \) is the predicted criterion variable given the \( X_1, \ldots, X_n \) cues and each \( b_{ei} \) coefficient represents the optimal weighting of cue \( i \). The correlation coefficient, \( R_e \), represents the linear predictability of the criterion event based on the cue set.

The right-hand side of the model can be represented in a similar manner:

\[ \hat{Y}_s = b_{s1}X_1 + b_{s2}X_2 + \ldots + b_{sn}X_n \]  

(2.3)

where \( \hat{Y}_s \) represents the prediction of the subject's response based on cues \( X_1, \ldots, X_n \) and \( b_{si} \) represents the extensiveness of the usage by the individual on cue \( X_i \). The correlation coefficient, \( R_s \) (called the "consistency index" by Dudycha and Naylor [1966]), represents the linear relationship between the individual's judgments and the given cue set.

As discussed by Ashton [1982b, p. 17] (see Figure 2), there are six correlation coefficients which may be calculated as indices of human judgment. Two important ones, \( R_e \) and \( R_s \), were described above. The other four are denoted as follows: \( r_a, r'_a, r_0, \) and \( G \). The so-called "achievement index" is represented by \( r_a \) and measures the individual's achievement in predicting the criterion variable. Related somewhat to that is \( r'_a \), which reflects the predictability of the criterion variable based on the individual's regression model.
The "optimality index" is represented by $r_0$ and measures the correspondence between the individual's judgments and those of the optimal regression model's prediction of the criterion variable. The "matching index" is represented by $G$ and reflects the correspondence between the individual's regression model and the optimal regression model of the environment.

The above discussion is presented from a linear multiple regression point of view. Analysis of variance (ANOVA) is also applicable to a lens model formulation. ANOVA is particularly useful in incorporating nonlinear or configural cue usage. "Main effects" represent the linear components of the regression model and "interactions" represent the configural components. A main effect merely indicates that the subject responds systematically to a particular cue while all other cues are held constant. A two-way interaction suggests that the way the individual responds to a given cue is conditional on the level of another cue. Three-way (and higher-order) interactions are more difficult to interpret (Keppel [1982]).

In the context of this study, no criterion variable exists. In many auditing contexts, the environment does not provide a "correct" outcome. (In rare instances, the courts may be viewed as providing such a determination on a case-by-case basis.) However, in most auditing situations there will not be a criterion variable for materiality. Hence, this study focuses on the right-hand side of the
focuses on consensus as a criterion. Since this study emphasizes comparisons among groups of auditors, rather than examining the judgments within individual auditors, consensus is of primary importance.

Ashton [1982b, pp. 124-152] reviewed a large number of studies in accounting which have adopted the lens model framework. The contexts included internal control evaluation, materiality judgments, audit report messages, bankruptcy prediction, stock recommendations and price predictions, functional fixation, and managerial decision making, among others. The materiality studies which Ashton described as lens studies (many of the judgment-capture experiments regarding auditors' materiality judgments discussed in detail above) were: Boatsman and Robertson [1974]; Moriarity and Barron [1976, 1979]; Hofstedt and Hughes [1977]; Firth [1979]; and Messier [1979]. In addition, although not included in Ashton's review of materiality-related lens studies, Schultz and Reckers [1981] and Krogstad, et al. [1984] would be appropriately viewed as applications of the lens model. Accordingly, there is ample empirical support for the lens model as a framework within which to explore individuals' judgments, in general, and auditors' materiality judgments, in particular. An advantage of the lens model is its applicability to diverse contexts.
object. Third, an attitude consists of positive or negative evaluations." (The same three characteristics were identified by Fishbein and Ajzen [1975, p. 6] in their definition of attitude.)

In addition, some issues still remain with respect to the measurement of attitudes. In their 1972 review article, Fishbein and Ajzen reported on over 750 articles published in 1968-1970 related to attitudes. They noted that several hundred different techniques or variations were used to measure attitudes in those articles.

Further, there is some uncertainty about the link between attitudes and behavior. Commenting on a renewed interest in research on attitude-behavior relationships, Ajzen and Fishbein [1977, p. 888] reported, "The emerging position seems to be that attitude is only one of many factors determining behavior." Furthermore, in promoting the "Theory of Mindfulness-Mindlessness," Palmerino, et al. [1984, p. 186], cautioned, "Attitudes, and contexts in general, have greater influence upon behavior when a person is mindless.... Laboratories may induce self-consciousness in our subjects and as such may be thought-provoking and habit-inhibiting." Consequently, attitudes which might be relevant to behavior at a "subconscious" level may be made irrelevant by raising the issues to a "conscious" level. Although attitudinal factors may only explain a portion of the observed behavioral differences, the investigation of differences in subjects' attitudes may, nevertheless, offer insights into differences in their behavior of interest.
These two factors are then weighted; the applicable weights may differ across individuals. The theory can be represented as follows:

\[ B \cdot I = (A_B)w_1 + (SN)w_2 \]  (2.5)

where \( B \) represents the behavior of interest; \( I \) represents the intention to perform the behavior \( B \); \( A_B \) represents the attitude toward performing behavior \( B \); \( SN \) represents the subjective norm associated with behavior \( B \); and \( w_1 \) and \( w_2 \) represent the individual's weights associated with the attitude and norm, respectively.

\( A_B \) was represented (2.4) as the summation of the product of the applicable beliefs and evaluations. The subjective norm, \( SN \), deals with the influence of the social environment on behavior, and can be represented as follows:

\[ SN = \sum_{i=1}^{n} b_i m_i \]  (2.6)

where \( b_i \) represents the normative belief that the behavior \( B \) should be performed in view of the reference group or individual \( i \); \( m_i \) represents the motivation to comply with the reference group or individual \( i \); and \( n \) represents the number of referents.

To the extent that the behavior and the intention to perform the behavior may be assumed to be highly related, then factors which influence the intention to perform the behavior will influence the behavior, too. Fishbein and Ajzen [1975, p. 369] consider three factors which may influence the relationship between intention and behavior: (1) the degree to which intention and behavior correspond in their levels of specificity; (2) the stability of intention; and
Thus, the psychology literature provides substantial empirical support for the constructs of the Fishbein attitude model. Recently, van der Plight and Eiser [1984, p. 165] commented on the Fishbein and Ajzen [1975] attitude model: "Since the introduction of this model research has shown a renewed interest in attitudinal phenomena...The theory of reasoned action has been applied to a wide variety of behaviors and a considerable amount of supporting evidence has accumulated in both experimental and naturalistic settings...."

In addition to the psychology literature, empirical support for the relevance of attitudinal differences in explaining behavioral differences can be found in the accounting literature. Kida [1980] is relevant to the issue of risk perception by auditors, and the consequent influence on their predispositions toward the alternative opinion choices. Kida studied auditors' decisions to issue going concern opinions to hypothetical clients relative to their classifications of the companies as troubled versus untroubled firms. He measured their attitudes toward issuing a qualified opinion to a financially troubled firm (a correct judgment) and toward issuing a qualified opinion to a financially untroubled firm (a Type I error); similarly, he measured their attitudes toward issuing an unqualified opinion to a troubled firm (a Type II error) and toward issuing an unqualified opinion to an untroubled firm (a correct judgment). Attitudes were measured first by semantic differential scales (which used bi-polar adjectives), and second by expectancy-valence (which
to behavior. In this study, behavior will consist of the reported materiality judgments of the participating auditors. As previously indicated, the attitude measure is a summation of the products of certain underlying beliefs ($b_i$'s) and evaluations ($e_i$'s). It is expected that auditors may identify different materiality thresholds because of different perceptions of the consequences (either differences in the perceived likelihoods of particular outcomes or in the perceived importance of those outcomes, or both) associated with auditors' materiality judgments. The expectancy-valence measure of attitude captures both of those dimensions by measuring $b_i$ and $e_i$ for each identified outcome, $i$.

Within the context of this study, the notion of "intention" would not appear to be directly relevant to auditors' materiality judgments. Perhaps one could incorporate the intention of an auditor to establish a "conservative" or a "liberal" materiality threshold. On a somewhat different level, one might consider the intention of an auditor to issue an unqualified opinion (versus a qualified opinion); this intention may be manifested by establishing higher (lower) materiality thresholds. However, it is believed that including intention in the study would not add any significant insights relative to focusing on behavior directly. In this study, the specified materiality thresholds are within the volitional control of each participating auditor. As a result, the participants should be able to
in all the other steps of their GAAS model. Although two firms had developed specific analytical approaches for quantifying the preliminary estimate of materiality and incorporating it into audit planning, they noted that five of the nation's twelve largest CPA firms did not have formal policies regarding preliminary estimates of materiality.

Kinney [1986] classified the Big Eight and several non-Big Eight firms as "structured," "intermediate," or "unstructured" based on their audit technologies.10 According to Kinney (p. 10), "Structured firms are more likely to argue that auditor consensus is desirable, that is, in a given situation different auditors should conduct similar audits and give the same opinion." Unstructured firms tend to leave more matters to the judgment of the individual auditors and to the practice offices. Morris and Nichols [1986, p.6] reported the specific measures of structure used by Kinney. Five individuals were asked to rate the degree of audit structure for twelve large CPA firms (including each of the Big Eight). Four of these individuals were members of the Auditing Standards Board who rated each of the firms; the other was someone within each firm who rated his or her own firm. The rating was a three-point scale (where "1" represented the least structured and "3" represented the most structured). Thus, the maximum possible score (summed across the five individuals) was fifteen and the minimum was five. The following scores were identified for the Big Eight firms, from most to least structured:
These include: Hofstedt and Hughes [1977]; Firth [1979]; Schultz and Reckers [1981]; Messier [1983]; and Krogstad, et al. [1984].

With one exception, all of the prior materiality-related research has focused on auditors' judgments at the evaluation stage. Only Moriarity and Barron [1979] explicitly examined materiality judgments at the planning stage; and that research occurred prior to the issuance of SAS Nos. 22 and 47, which required auditors to consider their "preliminary estimates of materiality" in audit planning. At that time, Moriarity and Barron reported that the task was unfamiliar to their eight subjects.

The dominant variable in all of the materiality studies has been the item in question as a percentage of net income. Other variables have not been consistently significant in the materiality judgments investigated.

Messier [1983] and Krogstad, et al. [1984], looked for experience effects associated with auditors' materiality judgments. The results have not been clear-cut. Using ANOVA, Messier found no mean difference among the experience groups for the materiality judgment; however, an experience effect was reported for the disclosure judgment. Krogstad, et al., reported that consensus was different for students relative to practitioner-subjects, but that consensus was not different for the audit seniors relative to the partners.

Woolsey [1954a], Dyer [1975], and Messier [1983] looked for systematic differences in the materiality judgments of auditors with
materiality thresholds related to attitudinal differences (as suggested by Crosby, et al. [1985]). Chapter III identifies the specific hypotheses to be examined, and discusses the research methodology to achieve the stated research objectives.
efficiency." Presumably, the ASB reasoned that the client's insistence would cause the auditor to investigate further and, therefore, reach the appropriate conclusion.

The ASB described another concept of risk, although not by name, in its footnote 1 of SAS No. 47. The SAS points out that this risk includes litigation and adverse publicity. "This exposure is present even though the auditor has performed his examination in accordance with generally accepted auditing standards and has reported appropriately on those financial statements" (AICPA [1983, note 1]). Brumfield, et al. [1983, p. 60], referred to this concept of risk as "business risk" which they defined as "... the probability that an auditor will suffer a loss or injury to his professional practice."

3. For a more detailed discussion of these pronouncements related to materiality from Canada, the United Kingdom, Australia, and New Zealand, see FASB [1975, paragraphs 90-103] and Selley [1984].

4. Holstrum and Messier [1982, p. 47] referred to a "... dearth of research concerning the use of the materiality concept in auditing ..." and observed that the focus of researchers and standard-setting bodies has been materiality in accounting. As described by Carmichael [1969], the two concepts are very much related, if not synonymous. Materiality in accounting deals with the fairness of the presentation of the financial statements; materiality in auditing deals with the design and execution of the audit approach to provide the auditor with the appropriate documentation and conclusions about the fairness of the financial statements. Selley [1984, pp. 10-11] summarizes his view of the distinction as follows:

... there is widespread agreement that materiality from the auditor's point of view at the final evaluation stage of the audit is the same as it is from the point of view of the preparers of the financial statements (so-called "accounting materiality"). However, SAS 47 has stated that the auditor for cost/benefit reasons may not in fact plan his procedures to detect some errors that would be qualitatively material. If this view prevails (as it is almost bound to do in the U.S.) we are again back to two materialities: "planning materiality" and "evaluation materiality."
\(x_{13}\) was the magnitude of stockholders' equity.

\(x_{14}\) was the same as \(x_2\), except that it was not divided by net income. (The idea here was to determine whether the absolute magnitude or relative magnitude or both were factors considered in materiality judgments.)

\(x_{15}\) was a dummy variable coded 1 if the firm reported a net loss and 0 otherwise.

\(x_{16}\) was a dummy variable coded 1 if the change turned net income into a loss or vice versa and 0 otherwise.

\(x_{17}\) was a dummy variable coded 1 if the CPA firm was a member of the "Big Eight" and 0 otherwise.

6. Moriarity and Barron cite Woolsey [1973b], and Pattillo and Siebel [1974] as evidence that auditors differ with respect to which variables they deem important in their materiality judgments. They cite Woolsey [1954a], Dyer [1975], and Pattillo [1975] as evidence that auditors differ in their materiality thresholds. Each of these studies was a survey and was, therefore, only amenable to limited statistical analysis.

7. In addition, Hofstedt and Hughes [1977] intended to partition their participants according to "risk-taking propensity," based on the Kogan-Wallach (twelve-item) Risk-taking Scale. However, they abandoned this analysis because (p. 384) "... (1) there was little difference between subjects [MBA students] and (2) the dimensionality of the measuring instrument is not clear...."

8. Messier's materiality judgment required his subjects to rank the materiality of a write-down of inventory (for each of 32 cases) on a seven-point scale. The disclosure task required the subjects to identify the probability that the write-down would be separately disclosed in the financial statements (i.e., a 101-point scale from "0" to "100").

It may not be surprising that Messier did not observe any mean difference among the experience groups (based on ANOVA) for the materiality judgments, since the sample consisted entirely of audit partners with varying years in practice. Within an internal control context, Ashton and Kramer [1980], which used students as subjects,
CHAPTER III -- HYPOTHESES AND METHODOLOGY

Introduction

The purpose of this chapter is to identify the research issues of interest, including the specific hypotheses to be investigated. In addition, this chapter describes the general methodology to be used, including the participants, the experimental task, the independent variables, the experimental design, and the planned analyses to achieve the stated research objectives and to test the specific hypotheses.

In general, the research issues of interest include better understanding how auditors may quantify their preliminary judgments about materiality, which variables they attend to, and whether a consensus appears to exist. In particular, since substantial individual differences are expected, the study concentrates on certain factors which may be associated systematically with such differences: level of responsibility, firm affiliation, and audit technology. As a secondary issue of interest, this research also considers the notion of differing risk attitudes to explore a possible reason why auditors may differ in their materiality thresholds.

Specifically, within the context of auditors' planning stage materiality judgments, this study focuses on the following: (1) the effect of level of responsibility on consensus;\(^1\) (2) the effect of
if a higher materiality threshold were adopted. Preliminary judgments about materiality which are "too high" may lead to the gathering of insufficient audit evidence and, as a result, audits which are not in accordance with Generally Accepted Auditing Standards (GAAS). The ASB (AICPA [1983, note 1]) cautioned against (presumably, establishing unduly high materiality thresholds and, thereby,) performing insufficiently extensive audit tests as a consequence of an auditor's perception of low exposure: "Even if an auditor assesses this exposure as low, he should not perform less extensive procedures than would otherwise be appropriate under generally accepted auditing standards." Conversely, preliminary judgments about materiality which are "too low" may lead to audit procedures which are "too extensive" and, thereby, audits which are not efficient (i.e., "over-auditing"). Auditors' materiality judgments are relevant to the issue of determining the amount of audit evidence that is obtained—to obtain a sufficient amount of evidence to support an auditor's opinion without obtaining more evidence than is necessary, since such evidence is costly.

The specific research hypotheses related to this study of auditors' planning stage materiality judgments will be described in the next section. The hypotheses are organized according to the research purposes introduced above. The primary research hypotheses deal with differences in consensus among auditors' planning stage materiality thresholds with respect to level of responsibility, firm
Hypotheses

Level of Responsibility

As described in Chapter II, it is not clear what effect experience has on auditors' materiality judgments, in general, and their preliminary judgments about materiality, in particular. Therefore, one issue in this research project is the effect of differing levels of responsibility (presumably related to experience) on the consensus of auditors' preliminary judgments about materiality. Consensus is examined in terms of two separate dimensions based on the recommendation of Gaumnitz, et al. [1982]--(1) location (regarding central tendency) and (2) variation (regarding dispersion). To address the effect of different levels of responsibility on auditors' preliminary judgments about materiality, the following hypotheses (in null form) are offered:

H1.1: Level of responsibility is unrelated to differences in location among auditors' planning stage materiality thresholds.

H1.2: Level of responsibility is unrelated to differences in variation among auditors' planning stage materiality thresholds.

Based on Krogstad, et al. [1984], which found no difference among seniors and partners, the three levels of responsibility are not expected to differ in their planning stage materiality judgments, either as to location or variation. However, Krogstad, et al. focused on an evaluation stage materiality judgment task, so their results may not extend to this planning stage task.
Audit Technology

A related issue involves audit technology. Cushing and Loebbecke [1983] and Kinney [1986] examined several characteristics as a basis for classifying firms according to audit structure (e.g., the extensiveness of sampling procedures). The classifications were consistent across the dimensions considered. That consistency, along with Cushing and Loebbecke's observation that there were conspicuous differences in the firms' treatments of planning stage materiality judgments, and the structured firms' emphasis on consensus, as described earlier by Kinney, provide the basis for believing that the notion of audit technology should be considered in this study. To the extent that differences are identified among the participating firms, systematic differences between the relatively structured and the relatively unstructured firms are of interest. Accordingly, the following hypotheses (in null form) will be examined:

H3.1: Audit technology is unrelated to differences in location among auditors' planning stage materiality thresholds.

H3.2: Audit technology is unrelated to differences in variation among auditors' planning stage materiality thresholds.

As discussed above, firm differences are expected. Further, based on Kinney [1986], differences as to location and/or variation, at the firm level, are expected to relate to differing audit technologies. In particular, the differences are expected to be
### True but unknown state:

<table>
<thead>
<tr>
<th>No Material Misstatements Exist</th>
<th>Material Misstatements Exist</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Type I Error</td>
<td>A Correct Judgment</td>
</tr>
<tr>
<td>{Cell #1}</td>
<td>{Cell #2}</td>
</tr>
<tr>
<td>A Correct Judgment</td>
<td>A Type II Error (&quot;Audit Risk&quot;)</td>
</tr>
<tr>
<td>{Cell #3}</td>
<td>{Cell #4}</td>
</tr>
</tbody>
</table>

**Action:**
- Concludes that there are material misstatements (qualified opinion -- "GAAP departure")
- Concludes that there are no material misstatements (unqualified opinion)

**Note:** The idea represented by the above matrix is that the auditor may have to render an opinion while uncertainties about some of the financial statement items remain. These uncertainties will not be resolved until a later point in time. It is hypothesized that auditors' individual differences with respect to their perceptions of the likelihoods and the consequences associated with these cells may be systematically related to the differences in their preliminary judgments about materiality.

**Figure 3**

(Implicit) Trade-offs Inherent in Auditors' Materiality Judgments
Independent Variables:
- NET SALES
- INCOME (LOSS) BEFORE TAXES
- TOTAL ASSETS

Intervening Variable:
- RISK ATTITUDE

Dependent Variables:
- PRELIMINARY JUDGMENT ABOUT MATERIALITY (INCOME STATEMENT)
- PRELIMINARY JUDGMENT ABOUT MATERIALITY (BALANCE SHEET)

Figure 4
Variables Used
Empirical Research Strategy

In order to achieve the study's research purposes, a number of characteristics should be associated with the chosen research strategy. In general, any study must be conducted with a reasonable cost (including a reasonable amount of time committed by any participants). Further, the obtrusiveness of the researcher should be minimized, in order to strengthen the validity of the study. Finally, to enhance generalizability (i.e., external validity), the study should be as realistic as possible.

Specifically, with respect to the stated research objectives, it is essential that the participants address the same scenarios. Since the study makes comparisons across the participants (in terms of level of responsibility, firm, and audit technology effects, and in terms of different risk attitudes), the participants must face identical cases for the comparisons to be meaningful.

In order to model the participants' judgments, several other attributes are desired. First, it is necessary that certain variables of interest be capable of being manipulated as independent variables. Second, and related to the first, is the need for multiple judgments by each participant. Third, the measurements must be objective and capable of relatively precise measurement. These qualities will make it possible to infer causality and enable the research to be replicated.
interest can be specified and their levels designated; and confounding factors can be controlled effectively. Multiple observations can be obtained across different levels of the independent variables. As a result, laboratory experiments are capable of replication. In addition, the cost of the research, including time demands on participants, can be modest.

The principal disadvantage of the laboratory experiment is its artificial setting. As a result of potentially weak external validity, it may not be meaningful to extrapolate the results of the experiment into realistic (natural) settings. In addition, Cook and Campbell [1979, pp. 50-80] describe a number of threats to the validity of research. Regarding the laboratory experiment strategy, a number of "artifacts" (including hypothesis-guessing by the participants, evaluation apprehension, and experimenter expectancies) may affect the results of an experiment.

In view of the stated research objectives, an experiment is the most appropriate approach for the present study. It will facilitate the desired comparisons to be made across subjects (in terms of level of responsibility, firms, audit technologies, and different risk attitudes), and permit the modeling of the participants' judgments. The required research costs, in particular the time demands on the participants, are reasonable. An incidental advantage of conducting an experiment is that it may facilitate any desired comparisons with
normally work. Permitting the participants to complete the experimental materials in their regular work environments also served to minimize the obtrusiveness of the researcher, since the subjects could complete the materials at their own convenience. Based on the other dimensions considered (i.e., cost, manipulation of independent variables, use of the same scenarios, multiple observations, and relatively objective measurements), an experimental approach is preferred.

**Subjects**

This study involves practicing auditors at the senior, manager, and partner levels. The "contact person" at each firm indicated that the seniors, managers, and partners make planning stage materiality judgments in practice. Based on representations by these individuals that staff-level subjects would not generally be involved with the real-world analogue of the experimental task, staff-level auditors were excluded from this experiment.

As described in Chapter II, there may be systematic differences in the judgments of auditors across Big Eight firms. In addition, there may be differences in the judgments of auditors with "structured" firms relative to auditors with "unstructured" firms. The Columbus, Ohio, offices of seven of the Big Eight firms were asked to participate; Arthur Young was not contacted due to the small size of their Columbus office. The following firms participated:
variables for which subjects were asked to provide their planning
stage materiality judgments at both the aggregate income statement and
balance sheet levels; (2) the attitude measurement questionnaire (see
Appendix B) which dealt with auditors' differing sensitivities to so-
called Type I and Type II errors; and (3) a survey or debriefing
questionnaire (see Appendix C) which solicited additional personal
information about each participant.

As described above, a laboratory experiment of audit judgment may
be criticized for its artificiality. However, that concern may be
mitigated to some extent by providing relatively realistic information
to the participants. The background information included the
principal items which SAS No. 22 suggested that auditors should
consider when involved in audit planning.9 Further, the case
materials were modeled from actual retail companies' annual reports,
and subjects were so informed.

Participants were asked to specify the dollar amounts which they
believed constituted the appropriate planning stage materiality
thresholds for the hypothetical client for each case. Specifically,
subjects were asked to make two materiality judgments for each case--
one representing their materiality threshold with respect to the
income statement; the other representing their materiality threshold
with respect to the balance sheet. (Figure 5 provides one page of the
document used to elicit subjects' income statement and balance sheet
materiality judgments.10) As described in Chapter II, SAS No. 47
(paragraph 12) indicates that a planning stage materiality judgment should be made at the aggregate financial statement level; the appropriate preliminary judgment about materiality would be the smallest materiality judgment applicable to any one of the financial statements. Consequently, it is appropriate to ask the participants to view the planning stage materiality judgment separately with respect to the income statement and the balance sheet. Subjects were permitted, if they desired, to refer to their firms' materials and other professional literature to assist them in their judgments.

As described in Chapter II, prior research has consistently reported net income (or income before taxes) to be the dominant variable in studies of materiality judgments at the evaluation stage. In those studies, balance sheet items (e.g., total assets) have tended to be unimportant variables. In addition, Zuber, et al. [1983, p. 43] asserted that planning stage materiality judgments are driven to a much greater extent by income statement rather than balance sheet considerations. Therefore, in this study, the income statement materiality judgments are deemed of primary interest. Nevertheless, comparable analyses will be conducted separately with respect to the income statement and balance sheet judgments, since so little is known about materiality judgments at the planning stage.

Subjects' risk attitudes were measured by their responses to a questionnaire (Appendix B), after their materiality judgments were elicited. The attitude measure used an expectancy-valence approach to
four scenarios above. As shown in Appendix B, the belief statements were represented by a seven-point scale where "1" represented "practically no chance that this would occur" and "7" represented "this almost certainly would occur." The evaluations were based on a five-point scale where "1" represented "neither bad nor good (indifferent)" and "5" represented "extremely bad" for unfavorable outcomes or "extremely good" for favorable outcomes.\textsuperscript{11}

A "survey" questionnaire (frequently referred to as a "debriefing" questionnaire in the accounting literature) was administered at the end of the experiment to accumulate certain demographic and other background information about each participant. (See Appendix C for the survey questionnaire.)

**Independent Variables**

As indicated above, this study was predicated on the expectation that there would be substantial individual differences among the participating auditors in their planning stage materiality thresholds. In order to gain further insights as to the nature of these differences, individual models for each subject's judgments were desired. The purpose of these individual models was to identify whether the structural form of the models differed among the subjects and whether any such differences in the structural forms of the models related to differences in their materiality thresholds. To make
for most auditors in quantifying their preliminary judgments about materiality with respect to the income statement. In addition to two levels of the variable which are relatively high (one may be viewed as "high" and the other as "medium"), one level was chosen at approximately the break-even level (since it was deemed of interest for future purposes to gather data about auditors' judgments as they approach a zero level of income). Finally, a loss situation was included of approximately the same absolute magnitude as the next to highest positive value; it was of interest for future purposes to gather data about whether the auditors' judgments were symmetrical between income and loss situations or whether the judgments were conditional on the positive or negative nature of the income (loss) before taxes. Only two levels of sales were included, since it was not expected that sales would be a significant variable for many of the participants. However, including two levels was sufficient to ascertain whether the level of sales was a significant variable in the auditors' judgments (either as a primary variable or as an interaction with one of the other variables). Three levels of total assets were chosen (representing "high," "medium," and "low") since it was expected to be the dominant variable with respect to the balance sheet materiality judgment. Using three levels made it possible to identify, for future purposes, whether the judgments with respect to the balance sheet were of a linear form. That was of interest since
Figure 6

Experimental Design

Income (loss) before taxes:

<table>
<thead>
<tr>
<th>IBT_1</th>
<th>IBT_2</th>
<th>IBT_3</th>
<th>IBT_4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(1,787,446)</td>
<td>$215,712</td>
<td>$1,429,976</td>
<td>$4,366,238</td>
</tr>
</tbody>
</table>

Net sales:

<table>
<thead>
<tr>
<th>S_1</th>
<th>S_2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$119,202,666</td>
<td>$205,790,182</td>
</tr>
</tbody>
</table>

Total assets:

<table>
<thead>
<tr>
<th>TA_1</th>
<th>TA_2</th>
<th>TA_3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$9,260,686</td>
<td>$35,388,746</td>
<td>$90,591,695</td>
</tr>
</tbody>
</table>
materiality judgments are of primary interest, and the results of the analyses of the income statement judgments are reported in Chapter IV. Since the balance sheet materiality judgments are of lesser interest, the results of the analyses are reported in Appendix D for the interested reader.

Conceivably, the subjects might differ systematically on many (or perhaps most) of the cases, without necessarily differing consistently on every case. Therefore, a single measure is desired as a basis for making overall inferences with respect to systematic differences among the subjects. This will avoid having to qualify the observations with respect to particular case numbers. This overall measure should capture the relative extremeness of each individual's judgments, as well as provide a reference point from which to measure variability within certain groupings of the participating auditors.

For all the subjects in the aggregate, the responses will be transformed to a distribution with mean zero and variance one, on a case-by-case basis. For a particular subject, an above average response for a given case would be represented by a positive score; an individual with a below average response on that case would be represented by a negative score. Each individual's standardized scores will be summed across the 24 unique cases to provide an overall measure of the extremeness of his or her materiality thresholds. These measures will be referred to as the "standardized materiality scores."
respective levels of the CPA firms may affect the priorities and concerns of the individual auditors.

Each subject can be viewed as a member of one of the three levels of responsibility, while simultaneously a member of one of the six participating firms, while simultaneously a member of one of two audit technologies. Consequently, subjects were nested within level of responsibility, within firm, within audit technology. In the context of this study, systematic differences among subjects' materiality thresholds might arise from two distinct sources—level of responsibility and firm affiliation (along with audit technology, which is related to firm affiliation). The effects associated with these different factors must be separated in order for the conclusions to be valid. The planned analyses reflect the need to separate the differences associated with level of responsibility from the differences associated with firm affiliation. The statistical tests will identify whether there are differences among the three levels of responsibility, first as to location and second as to variation.

Location

The first analyses to be conducted are concerned with differences among the three levels of responsibility with respect to location (i.e., central tendency). A one-way ANOVA involving levels of responsibility is not appropriate because such an analysis ignores the influence of firm affiliation on subjects' responses while considering
where:

Each $A_1, A_2, A_3, \ldots, C_6$ represents 4 subjects;
each $A_s, C_s, A_u, C_u$ represents 12 subjects;
each $A, C$ represents 24 subjects;
each $T_1, \ldots, T_6$ represents 12 subjects;
both $T_5$ and $T_u$ represent 36 subjects;
$T$ represents 72 subjects.

**Figure 7**

Notation Regarding Comparisons To Be Made
dealing with level of responsibility (first among the structured firms, and second among the unstructured firms).

**Variation**

Similar comparisons will be conducted to identify any differences among the levels of responsibility with respect to variation. Overall tests for homogeneity of variance cannot be appropriately conducted in the aggregate among the levels of responsibility, since that would ignore the potential firm effect as a source of variation. Accordingly, the tests for homogeneity of variance will be applied on a firm-by-firm basis.

The sums of squares, related to variance, will be calculated for each level of responsibility, separately for each firm, based on the standardized materiality scores. Cochran's test for homogeneity of variance (Kirk [1982, p. 78]) will be applied, firm-by-firm, to the sums of squares associated with the seniors, managers, and partners. The p-values will be calculated with respect to level of responsibility for each firm. As described above, Fisher's method for combining tests will be used to combine the p-values for the six firms into an overall test of H1.2 (regarding level of responsibility effects with respect to variation).

A supplemental comparison examines differences among the three levels of responsibility separately within each of the two audit technologies. With respect to differences in variation, Cochran's
Table 4
Analyses of Level of Responsibility Effects

<table>
<thead>
<tr>
<th>Comparisons To Be Made (Notation per Figure 7)</th>
<th>Data To Be Used</th>
<th>Primary Method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I. Location:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Level of responsibility, by firm affiliation</td>
<td>Materiality scores</td>
<td>One-way ANOVA for each firm</td>
</tr>
<tr>
<td>A&lt;sub&gt;1&lt;/sub&gt;, B&lt;sub&gt;1&lt;/sub&gt;, C&lt;sub&gt;1&lt;/sub&gt;</td>
<td>standardized across all 72 subjects,</td>
<td></td>
</tr>
<tr>
<td>A&lt;sub&gt;2&lt;/sub&gt;, B&lt;sub&gt;2&lt;/sub&gt;, C&lt;sub&gt;2&lt;/sub&gt;</td>
<td>partitioned by firm</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A&lt;sub&gt;6&lt;/sub&gt;, B&lt;sub&gt;6&lt;/sub&gt;, C&lt;sub&gt;6&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Overall (H1.1)</td>
<td>Same as above</td>
<td>P-values for firms combined by Fisher's method</td>
</tr>
<tr>
<td>Same as above</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Level of responsibility, by audit technology</td>
<td>Same as above,</td>
<td>One-way ANOVA for each audit technology</td>
</tr>
<tr>
<td>A&lt;sub&gt;S&lt;/sub&gt;, B&lt;sub&gt;S&lt;/sub&gt;, C&lt;sub&gt;S&lt;/sub&gt;</td>
<td>but partitioned by audit technology</td>
<td></td>
</tr>
<tr>
<td>A&lt;sub&gt;U&lt;/sub&gt;, B&lt;sub&gt;U&lt;/sub&gt;, C&lt;sub&gt;U&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>II. Variation:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Level of responsibility, Sums of squares (SS's)</td>
<td>Same as I(A) above</td>
<td>Cochran's test for homogeneity of variance</td>
</tr>
<tr>
<td>by firm affiliation</td>
<td>for each level of responsibility for each firm</td>
<td></td>
</tr>
<tr>
<td>Same as I(A) above</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Overall (H1.2)</td>
<td>Same as II(A) above</td>
<td>P-values for firms combined by Fisher's method</td>
</tr>
<tr>
<td>Same as I(A) above</td>
<td>Same as II(A) above</td>
<td></td>
</tr>
<tr>
<td>C. Level of responsibility, SS's associated by audit technology</td>
<td>Same as I(C) above</td>
<td>Cochran's test for homogeneity of variance</td>
</tr>
<tr>
<td>with each level of responsibility for each audit technology</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Location**

H2.1 (regarding differences at the firm level with respect to location) will be tested by running a one-way ANOVA of the standardized materiality scores. If firm affiliation is found to be significant, Tukey's pairwise comparison test will be used to identify which firms differ from one another.\(^\text{19}\)

A supplemental comparison explores whether firm effects are present at each (or any) of the levels of responsibility. A one-way ANOVA will be run separately for each level of responsibility. In each instance that a significant firm difference is detected, Tukey's pairwise comparison test will be used to identify which firms differ.

Similarly, another supplemental comparison examines firm differences, within each audit technology group. (There are three firms in each of the two audit technology groups.) A one-way ANOVA will be run separately for each of the two audit technology groups. Again, for any significant result, Tukey's pairwise comparison test will be used to identify which of the firms differ.\(^\text{20}\)

**Variation**

To test H2.2 (regarding firm differences with respect to variation), Cochran's test for homogeneity of variance will be applied to the sums of squares associated with each firm.\(^\text{21}\) This test will indicate whether the variances among the firms are the same or whether at least one of the firms is significantly more variable than at least
Table 5
Analyses of Firm Effects

<table>
<thead>
<tr>
<th>Comparisons To Be Made (Notation per Figure 7)</th>
<th>Data To Be Used</th>
<th>Primary Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Location:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Overall (H2.1)</td>
<td>Materiality scores standardized across all 72 subjects</td>
<td>One-way ANOVA across all 6 firms</td>
</tr>
<tr>
<td>(T_1, T_2, \ldots, T_6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Firm affiliation, by level of responsibility</td>
<td>Same as above, but partitioned by level of responsibility</td>
<td>One-way ANOVA for each level</td>
</tr>
<tr>
<td>(A_1, A_2, \ldots, A_6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(B_1, B_2, \ldots, B_6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C_1, C_2, \ldots, C_6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Firm affiliation, by audit technology</td>
<td>Same as above, but partitioned by audit technology</td>
<td>One-way ANOVA for each audit technology</td>
</tr>
<tr>
<td>(T_1, T_2, T_3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(T_4, T_5, T_6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>II. Variation:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Overall (H2.2)</td>
<td>Sums of squares (SS's) associated with each firm</td>
<td>Cochran's test for homogeneity of variance</td>
</tr>
<tr>
<td>Same as I(A) above</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Firm affiliation by level of responsibility</td>
<td>SS's associated with each level of responsibility within each firm</td>
<td>Cochran's test for homogeneity of variance</td>
</tr>
<tr>
<td>Same as I(B) above</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Firm affiliation by audit technology</td>
<td>SS's associated with each firm, partitioned by audit technology</td>
<td>Cochran's test for homogeneity of variance</td>
</tr>
<tr>
<td>Same as I(C) above</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Firms #1-3 constitute the structured firms, whereas firms #4-6 constitute the unstructured firms.

Each level of responsibility will be examined separately to identify whether the structured and unstructured firms differ. Therefore, three separate one-way ANOVAs dealing with audit technology will be conducted. This supplemental comparison is intended to identify whether the different audit technologies have an effect on any or all of the levels of responsibility.

Variation

To test H3.2 (regarding audit technology effects with respect to variation), Cochran's test for homogeneity of variance will be applied to the sums of squares associated with the structured and unstructured firms. A supplemental analysis explores each level of responsibility separately for differences in variability associated with audit technology. The sums of squares for each level of responsibility will be calculated within each audit technology group. Cochran's test for homogeneity of variance will be applied to those sums of squares, on a level-by-level basis. This analysis will provide insights as to the consistency of any overall audit technology effect across the three levels of responsibility. Table 6 summarizes the primary analyses related to audit technology effects.
Risk Attitude

The notion of risk attitude is concerned with auditors' relative sensitivities to Type I errors (i.e., issuing a qualified opinion on financial statements which do not contain material misstatements) versus Type II errors (i.e., issuing an unqualified opinion on financial statements which do contain material misstatements). As hypothesized above, a difference among auditors in their attitudes about the implicit tradeoff of Type I and Type II errors (H4.1) has implications with respect to their relative materiality thresholds (H4.2). Subjects' attitudes toward Type I and Type II errors will be measured by means of a questionnaire which adopts an expectancy-valence approach. As described above and as represented by Figure 3, the attitude measures are organized according to four cells—representing each combination of a simple action set (issue a qualified opinion; issue an unqualified opinion) and a simple state set (the financial statements do not contain any material misstatements; the financial statements do contain material misstatements).

Cells #1 and #4 of Figure 3 represent Type I and Type II errors, respectively. Each cell specifies seven outcomes that might be associated with the particular cell. (Table 7 identifies the seven items specified for each cell.) The questionnaire elicits subjects' views of the likelihood ("belief") associated with each outcome for
each cell along with the importance ("evaluation") of the outcome. The product of the belief and evaluation, termed the "expectancy-valence," will be calculated for each outcome within each cell. The expectancy-valences for the outcomes will be summed within the cells for each participant to obtain a measure of each individual's concern for Type I (i.e., cell #1) and Type II (i.e., cell #4) errors in an absolute sense.

The relevant notion here is the relative sensitivity of an individual to Type I and Type II errors. In other words, it is the individual's tradeoff of Type I versus Type II errors that is of interest. Thus, a single measure of each subject's risk attitude will be calculated as the expectancy-valence measure associated with Type II errors (i.e., cell #4) divided by the measure associated with Type I errors (i.e., cell #1). Individuals who are relatively more concerned with, for example, litigation by financial statement users (associated with Type II errors) would have larger risk attitude measures than individuals who are relatively more concerned with, for example, practice development (associated with Type I errors).

To construct a reasonably powerful test, subjects will be partitioned into the upper, middle, and lower one-thirds, based on the extremeness of their standardized materiality scores. To test H4.1, the subjects comprising the upper one-third will be compared to the subjects comprising the lower one-third for differences in their risk attitudes; a one-way ANOVA will be run on the risk attitude measures.
standardized materiality scores for the subjects in the aggregate, using the parametric Pearson product-moment correlation. As suggested by the above discussion, a negative correlation is expected.

**Additional Analyses**

Finally, five additional types of analyses will be conducted. First, individual ANOVA models will be constructed for each subject's judgments. The purpose is to identify any apparent differences among the relevant groupings of subjects (in terms of level of responsibility, firm affiliation, and/or audit technology) with respect to the structural forms of their models. Second, business risk will be examined as a possible confounding factor. Third, test-retest consistency will be examined based on the four repeated cases. Fourth, order effects will be investigated for differences as to location and/or variation between the two orders of case presentation. Fifth, several (ad hoc) comparisons will be made to identify whether the income statement materiality thresholds are less than the balance sheet materiality thresholds, on average as expected (in terms of level of responsibility, firm affiliation, and audit technology).

**Individual ANOVA Models**

Individual ANOVA models will be calculated to model each participant's planning stage materiality judgments, separately for the
judgments, the examination of business risk is conducted with respect to level of responsibility, firm affiliation, and audit technology. To ascertain whether subjects differ in their business risk assessments, separate one-way ANOVAs will be run on a case-by-case basis with respect to level of responsibility, firm affiliation, and audit technology.

**Test-retest Consistency**

Test-retest consistency will be examined by repeating four of the cases at the end of the judgment task; the judgments on these repeated cases will be compared with the original judgments for the same cases. If subjects are found to be inconsistent between the original and repeated cases, then their original responses may be suspect. Such results would suggest that subjects were either unwilling/unable to provide meaningful responses to the experimental materials or that they were inattentive to the experimental task.

The analyses with respect to test-retest consistency will be performed on a case-by-case-basis for the four repeated cases. First, the responses to the repeated cases will be correlated with those on the original cases to ascertain the degree to which the responses move in the same direction. This correlation will merely identify whether relatively large responses on the original cases are paired with relatively large responses on the repeated cases. Second, for each of the repeated cases, a paired t-test will be calculated to ascertain
orders were systematically varied over level of responsibility, firm affiliation, and audit technology. Half of the subjects (at each level within each firm) received order #1 and the others received order #2.

A one-way ANOVA will be used to identify any difference as to location between the two orders of presentation, based on the standardized scores of the subjects receiving each order. Cochran's test for homogeneity of variance will be used to identify any difference as to variation between the two orders of presentation, based on the sums of squares associated with the subjects receiving each order.²⁷

Income Statement Versus Balance Sheet Judgments

It was stated above that the income statement materiality judgments are of primary interest in this study. The balance sheet materiality judgments were included because SAS No. 47 indicates that such judgments may be made at the aggregate balance sheet level. Consequently, the balance sheet judgments were included for exploratory purposes. The SAS indicates that, as a practical matter, a single "preliminary judgment about materiality" should be determined. It is identified as the smallest of the amounts material to any one of the financial statements. Prior research has consistently reported the dominance of net income (or income before taxes) in auditors' evaluation stage materiality judgments. In
financial statement level. The income statement judgments are expected to be less than the balance sheet judgments (hence, they will drive the auditors' "preliminary judgments about materiality") for most of the auditors, for most of the cases. Since this dissertation emphasizes systematic differences in terms of level of responsibility, firm affiliation, and audit technology, these ad hoc comparisons are made at those levels.

Underlying Statistical Assumptions

This section identifies the principal assumptions underlying these statistical methods to be used in the data analysis and discusses the reasons that those assumptions are not believed to constitute problems for the validity of the planned analyses. As described above, the judgments of the participants will be examined for systematic differences with respect to location and variation. The primary statistical method to be used to identify differences as to location is analysis of variance (ANOVA), a parametric method. Tukey's all pairwise multiple comparison test will be used to identify which levels of responsibility and/or which firms differ, when the related ANOVA is significant.

There are three principal assumptions underlying analysis of variance: (1) normally distributed populations; (2) homogeneous error variances associated with the populations; and (3) independent error
smallest variances), especially when the observed F-value falls near the critical F-value. Once again, this assumption should not be consequential, since the cell sizes will be equal by design for all comparisons. In addition, as noted, when the parametric tests are significant with respect to the primary hypotheses (i.e., H1.1, H1.2, H2.1, H2.2, H3.1, and H3.2), nonparametric Wilcoxon rank sum tests will be used. These nonparametric tests will provide evidence on the importance of the homogeneity of variance assumption.

The third assumption involves the independence of the error components of the ANOVA model. The error terms are assumed to be independent, both within and between comparison groups. Keppel (p.87) paraphrases this assumption: "...this is just another way of saying that systematic biases must not be present in the assignment of subjects to conditions." Independence within comparison groups requires that each observation be unrelated to all other observations. Independence between comparison groups requires that there be no confounding effects which would distort the inferences with respect to the relationship between the independent and dependent variables. Random selection of subjects is frequently used to achieve independence.

The standardized scores for each subject should be independent measures of each individual's relative extremeness in the materiality judgments examined. One way in which the independence of the data could be jeopardized would be by subjects working on the experiment
associated with the F-test are as follows: observations are random samples from normally distributed populations; the numerator and denominator of the F ratio are estimates of the same population variance; and the numerator and denominator of the F ratio are independent (Kirk [1982, pp. 56-57]). Kirk (p. 77) points out that the F-test is quite robust with respect to violations of the normality assumption; in addition, it is also robust with respect to violations of the heterogeneity assumption when the cell sizes are equal. (By design, the cell sizes for all comparisons are equal.) Further, nonparametric Wilcoxon rank sum tests will be used to cross-check the results when significance regarding the primary hypotheses is established by the parametric methods.

**Administration of the Experiment**

During late autumn of 1985, a "contact person" (a partner) at seven of the Big Eight firms' Columbus offices was asked for assistance in conducting the experiment. Arthur Young was not asked to participate due to the small size of its Columbus office. Ernst & Whinney declined to participate because they were too busy at the time. The other six Big Eight firms agreed to participate and help in any way that they could.

Prior to conducting a pilot study, the contact person at each of the firms expressing a willingness to participate in the study was
between-subjects variables (one distinguishing public versus private companies; and the other representing high client-paid fees versus low client-paid fees). The nature of the materiality judgments under study was apparently so "noisy" that the between-subjects variables were likely to contribute additional noise without offering much explanatory power. Thus, the focus of the study became the differences among auditors related to differences in their materiality judgments, rather than focusing on situational factors and their impact on auditors' materiality judgments. The further study of certain situational factors will be among the topics for future research.

In May of 1986 the experimental materials were delivered to the contact person at each firm to oversee distribution to the participants. (In only one case were the materials provided directly to the participants by the researcher. That involved Price Waterhouse, which had already arranged a staff meeting for the entire Columbus office and it was convenient for the firm for the researcher to distribute the materials directly to the chosen participants.) The completed booklets were returned to a designated person within each firm; the researcher then picked up these materials periodically as they became available.

Although deadlines were agreed upon during June, a number of individuals did not complete the experiment within the planned time-frame and some follow-up was necessary. The booklets were kept track
CHAPTER III FOOTNOTES

1. As referenced in Chapter II, Ashton [1982a, p. 74] notes that audit decision making has been evaluated with respect to six criteria: (1) accuracy; (2) normativeness; (3) consistency with professional pronouncements; (4) consensus; (5) self-insight; and (6) judgment stability. The first three criteria are not applicable to this study. There is no available criterion variable by which to assess optimality; and, due to the generality of the auditing standards (i.e., SAS No. 47), inconsistencies with the pronouncements are not applicable to an empirical study such as this. (See Cushing and Loebbecke [1983] for an evaluation of 12 national firms' policies relative to the SAS's.)

Ashton notes that consensus, self-insight, and stability are usually used when accuracy and normativeness cannot be used. Since this study focuses on comparisons among groups of auditors, rather than within individual auditors, consensus is of primary importance. Prior research (e.g., Messier [1983]; and Krogstad, et al. [1984]) has reported relatively high self-insight and stability on materiality tasks.

2. In addition to preceding the final materiality judgment, one could argue that quantifying the preliminary judgment about materiality is a more complicated task. That is, it may be more difficult to specify a dollar amount as a basis for audit planning (i.e., quantifying the preliminary judgment about materiality) than, for example, to make the dichotomous judgment that a specific item under consideration warrants separate disclosure, or that an account warrants adjustment.

On the other hand, one might argue that the evaluation stage judgment is more complicated. SAS No. 47 indicates that the evaluation stage judgment requires consideration of both quantitative and qualitative factors; the planning stage judgment primarily involves just the quantitative factors. Thus, one could argue that the judgment at the evaluation stage is more complicated in the sense of involving a larger set of variables which might enter into the judgment. As suggested by Woolsey [1973], there may be a "border zone" such that below some threshold there is a presumption of "immateriality;" above some other threshold there is a presumption of "materiality." Intuitively, the endpoints of this border zone would apparently hinge on quantitative factors. Within the border zone, the determination of materiality may hinge on the qualitative factors. Only within this "grey area" would the evaluation stage materiality
The partition of the firms into two groups—i.e., highly structured ("structured") and relatively less structured ("unstructured")—relies on the ordering identified by Kinney [1986]. The set of highly structured firms was comprised of Deloitte, Haskins, & Sells; Peat, Marwick, Mitchell; and Touche Ross.

For purposes of this study, Arthur Andersen was combined with Coopers & Lybrand and Price Waterhouse to comprise the set of relatively less structured firms. Kinney classified Arthur Andersen as an "intermediate" firm. Review of the firms' audit manuals with respect to planning stage materiality judgments suggested that these three firms were similar in their treatment of planning stage materiality judgments, at the time the experiment was conducted.

Consequently, the classification of Arthur Andersen as "unstructured" was slightly subjective and reflected a somewhat arbitrary cut-off along the continuum of unstructured to structured audit technology. In general, there can be no guarantee that, within either of the two audit technology groups, the firms reflect identical degrees of structure.

8. The information contained in Appendix A specifies the role as "partner." The materials provided to each subject specified the appropriate role in terms of "senior," "manager," and "partner." With the exception of identifying the role conditional on each subject's level of responsibility, the information provided to each subject was identical.

9. SAS No. 22 (AICPA [1978, paragraph 3]) identifies a number of factors that the auditor should consider in audit planning:
   a. the entity's business and industry environments (including competitive conditions);
   b. the entity's accounting policies;
   c. whether reliance on internal controls is anticipated;
   d. "preliminary judgments about materiality levels ..." (i.e., the dependent variables of interest in this study);
   e. statement items likely to require adjustment;
   f. related party transactions;
   g. the nature of the reports expected to be issued.

   These factors, among others, were explicitly addressed in the background information regarding the hypothetical client (Appendix A).

10. As indicated by Figure 5 and as explained in Appendix A (page 1), subjects were also asked to identify the business risk implied by the given financial statement amounts. (Business risk is defined on page 1 of Appendix A, and is often referred to as "exposure" in practice.) Risk was assessed on a seven-point scale from very low to very high.
15. According to Keppel [1982, p. 165], Tukey's multiple comparison test is generally preferred to a number of available alternative tests for the purpose of making all pairwise comparisons among means. Its advantages include relative simplicity, favorable power characteristics, and appropriate control of alpha in view of the number of comparisons made.

16. As indicated by Kirk [1982, pp. 57-58], the total variance for np scores in an experiment is given by:

$$\frac{\sum_{i=1}^{n} \sum_{j=1}^{p} (Y_{ij} - \bar{Y})^2}{np}$$

The numerator above is referred to as "the sum of squares." When the cell sizes (i.e., p) for n comparison groups are equal, the ratio of their variances is equal to the ratio of their sums of squares.

17. The available tables (e.g., Kirk [1982, p. 829]) provide the critical values for Cochran's test for homogeneity of variance at the .05 and .01 levels. However, these tables are not amenable to identifying particular p-values. In order to use Fisher's method for combining tests, the specific p-values must be derived. Cochran [1941] demonstrates how such p-values can be calculated. For three variances (i.e., for the seniors, managers, and partners), the following formula calculates the p-value for one of the groups:

$$p = 3 \left[ 1 - 6.5625 \left( 2 \frac{z^{3/2}}{3} - 4 \frac{z^{5/2}}{5} + 2 \frac{z^{7/2}}{7} \right) \right]$$

where $$z = \max_{i \leq j \leq 3} \frac{s_i^2}{\sum_{j=1}^{3} s_j^2}$$

This formula is valid for $$z \geq .5$$; for $$z < .5$$, using $$z = .5$$ gives a lower bound for the p-value.

This approach will be used to infer the specific p-value from Cochran's test for homogeneity of variance for each level of responsibility. Then, Fisher's method will be used to combine the separate p-values into a single test of significance.

18. If significant effects for level of responsibility are found, then firm differences will be evaluated, separately, for each level of responsibility. Fisher's method for combined experiments would be used to combine the p-values associated with each level of responsibility into an overall test of significance regarding firm affiliation. These procedures would be similar to those previously described with respect to level of responsibility in terms of location and variation.
two firms' sums of squares (or variances, as explained in footnote #16 above) constitutes an F-test. This represents a specific condition of the more general Hartley's test for homogeneity of variance, since, in this case, only two variances are involved in the comparison (Kirk [1982, pp. 54 and 78]). (In view of the number of such comparisons, the level of significance will be established conservatively--i.e., approximately .05/15.)

Of the resulting fifteen pairwise comparisons, three are comparisons between structured firms, three are comparisons between unstructured firms, and nine are comparisons between one structured and one unstructured firm. Consequently, these F-tests also relate to the analysis of firm differences within audit technology groups and to the analysis of audit technology effects (i.e., regarding differences between firms having different audit technologies).

23. If significant experience-related effects are found, then audit technology differences will be evaluated level-by-level for each level of responsibility. Fisher's method for combined experiments would be used to combine the p-values associated with each level of responsibility into an overall test of significance regarding audit technology. These procedures would be similar to those previously described with respect to level of responsibility regarding location and variation.

24. In the event that the ANOVA detects a significant audit technology difference, the nonparametric Wilcoxon rank sum test (Hollander and Wolfe [1973, pp. 27-28] will be used to cross-check the parametric results. (There would be 36 subjects in each technology group and, hence, the large sample approximation would be used.) Corroboration would be viewed as strong evidence that the results of the ANOVA do not depend upon the distributional assumptions of ANOVA.

25. If audit technology differences in variability are shown using the parametric Cochran's test for homogeneity of variance, a nonparametric Wilcoxon rank sum test will be used as a cross-check. Corroboration using a nonparametric approach would demonstrate that the assumptions underlying the parametric test are not driving the results.

   Specifically, the variance within each level of responsibility for each firm will be calculated (thus, there will be 9 observations for each audit technology). The Wilcoxon rank sum test will then be used to identify whether there is a difference in variability between the two technologies.

26. Additional one-way ANOVA's will be run with respect to subjects' risk attitudes (including the individual belief and evaluation
CHAPTER IV -- RESULTS OF ANALYSES

Introduction

As described in Chapter III, each participant was asked to make two separate materiality judgments (one with respect to the income statement and the other with respect to the balance sheet) for each of the 24 unique combinations of the three independent variables—total assets, net sales, and income (loss) before taxes. Separate analyses were conducted on the income statement materiality judgments and on the balance sheet materiality judgments.

As discussed in Chapter II, prior research (substantially directed at the evaluation stage) has consistently reported net income (or income before taxes) to be the dominant consideration in the judgments studied. Further, in practice the auditor's planning stage materiality judgment is believed to be driven primarily by income statement considerations (e.g., see Zuber, et al. [1983, p. 43]). Accordingly, the income statement materiality judgment is deemed to be the dependent variable of principal interest in this study, and this chapter reports the analyses only with respect to these income statement judgments. However, since SAS No. 47 (paragraph 12) implies that the planning stage materiality judgment may be made appropriately from a balance sheet perspective, similar analyses were performed on the balance sheet materiality judgments for exploratory purposes. The
(2) firm affiliation, and (3) audit technology (i.e., "structure"), related to firm affiliation. These comparisons emphasize consensus. This investigation of consensus involves two dimensions—location (regarding central tendency) and variation (regarding dispersion).

Effects associated with level of responsibility were examined among the seniors, managers, and partners. No significant differences were identified either with respect to location or variation.

Differences as to location and variation were examined among the six participating Big Eight CPA firms. Significant firm differences were identified among the participants both with respect to location and variation.

Further analyses were conducted to identify any differences as to location and variation related to the firms' differing audit technologies. The firms were partitioned into two groups—the structured firms and the unstructured firms. Significant differences were identified between the structured and unstructured firms both with respect to location and variation. The participants affiliated with structured firms had higher materiality thresholds, but greater variability, relative to the participants from the unstructured firms.

In addition, this chapter includes discussions of the results of the analyses with respect to risk attitudes. Risk attitude was viewed as a potential intervening variable affecting auditors' materiality judgments. Auditors who were relatively more sensitive to Type I errors (e.g., practice development considerations) were expected to
judgments, rather than the balance sheet judgments. These *ex post* comparisons underscored the appropriateness of that emphasis.

This chapter concludes with comments about the statistical assumptions underlying the principal analyses. Reasons are given as to why those assumptions are not viewed as problematic.

**Level of Responsibility Effects**

As discussed in Chapters II and III, it is unclear whether auditors' preliminary judgments about materiality differ systematically among auditors of differing levels of responsibility (presumably related to experience). Hypotheses were stated with respect to overall level of responsibility effects, separately in terms of location and variation. H1.1 deals with differences as to location; and H1.2 deals with differences as to variation.

These analyses distinguish the responses of the seniors, managers, and partners. The demographic information that was obtained provided the following profile of the subjects' years of experience:

<table>
<thead>
<tr>
<th>Level of Responsibility</th>
<th>Range</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seniors</td>
<td>2.0-6 years</td>
<td>3.9</td>
<td>1.3</td>
</tr>
<tr>
<td>Managers</td>
<td>4.5-14 years</td>
<td>7.9</td>
<td>2.1</td>
</tr>
<tr>
<td>Partners</td>
<td>9.5-40 years</td>
<td>17.6</td>
<td>6.3</td>
</tr>
</tbody>
</table>

Similar profiles were constructed separately for each firm and are reported in Table 8. Years of experience overlap with level of
responsibility in some instances, but are obviously correlated with level of responsibility. (As noted in Chapter III, the overall correlation coefficient between years of experience and level of responsibility was .84.)

As previously indicated, this dissertation investigates two distinct potential sources of systematic differences among the subjects (experience-related effects and firm-related effects.) Level of responsibility and firm effects were separated by analyzing level of responsibility effects within each firm; the results for each firm were then combined into an overall test of significance with respect to level of responsibility. In terms of Figure 7, the following comparisons were made: $A_1$, $B_1$, and $C_1$; ...; and $A_6$, $B_6$, and $C_6$.

In order to further explore any possible differences associated with level of responsibility, an additional type of comparison was made, and the results are also reported below. Each of the three levels of responsibility was compared within the two audit technologies (structured and unstructured). In terms of Figure 7, the following comparisons were made: $A_s$, $B_s$, and $C_s$; and $A_u$, $B_u$, and $C_u$. As indicated, formal hypotheses were not stated with respect to this supplemental comparison, since the purpose was to glean possible insights regarding effects associated with level of responsibility.

As described above, in order to obtain a single, independent measure of the extremeness of each subject's materiality judgments and to mitigate the substantial heterogeneity of variance with respect to
the individual cases, the responses for the seniors, managers, and partners were standardized on a case-by-case basis. For each case, the materiality judgments were transformed to a distribution with mean zero and variance one, across the 72 subjects. To get a measure of the extremeness of the individuals' judgments, each subject's standardized scores were summed across the 24 cases. Figure 8 plots these standardized materiality scores by level of responsibility (where levels 2, 3, and 4 identify the seniors, managers, and partners, respectively). These analyses are discussed first with respect to location and second with respect to variation.

**Location**

H1.1 deals with differences as to location among the seniors, managers, and partners. (Differences in location with respect to their standardized materiality scores would be interpreted as differences in their mean materiality thresholds.) As noted above, Figure 8 plots the standardized materiality scores by level of responsibility. Examination of that plot does not indicate any obvious differences in the means among the seniors, managers, and partners.

Six separate one-way ANOVAs were conducted to identify whether there were any detectable differences as to location among the seniors, managers, and partners of each participating firm. As summarized in Table 9, level of responsibility was significant
Table 9

Standardized Materiality Scores and Averages
By Level of Responsibility and By Firm Affiliation

<table>
<thead>
<tr>
<th>Firm #1</th>
<th>Seniors</th>
<th>Managers</th>
<th>Partners</th>
<th>Average</th>
<th>p-value *</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17.37</td>
<td>32.60</td>
<td>24.39</td>
<td>24.79</td>
<td>.77</td>
</tr>
<tr>
<td>2</td>
<td>30.41</td>
<td>-2.57</td>
<td>-10.47</td>
<td>5.79</td>
<td>.01</td>
</tr>
<tr>
<td>3</td>
<td>11.75</td>
<td>3.93</td>
<td>-5.30</td>
<td>3.46</td>
<td>.16</td>
</tr>
<tr>
<td>4</td>
<td>-13.94</td>
<td>-10.77</td>
<td>-14.15</td>
<td>-12.95</td>
<td>.62</td>
</tr>
<tr>
<td>5</td>
<td>-12.75</td>
<td>-8.88</td>
<td>-10.98</td>
<td>-10.87</td>
<td>.39</td>
</tr>
<tr>
<td>6</td>
<td>-15.68</td>
<td>-10.05</td>
<td>-4.92</td>
<td>-10.22</td>
<td>.06</td>
</tr>
<tr>
<td>Average</td>
<td>2.86</td>
<td>.71</td>
<td>-3.57</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p-value based on one-way ANOVA for level of responsibility calculated on a firm-by-firm basis.
significant for either the structured or the unstructured firms. The p-value was approximately .20 for the structured firms and .08 for the unstructured firms. The p-value for the unstructured firms might be viewed as borderline significant. Intuitively, one might expect experience-related factors to be more influential within the unstructured firms (relative to the structured firms, which have developed more explicit guidelines to provide more direction to their auditors at all levels of responsibility). For the structured firms, the means were approximately 19.8, 11.3, and 2.9 for the seniors, managers, and partners, respectively. For the unstructured firms, the means were approximately -14.1, -9.9, and -10.0 for the seniors, managers, and partners, respectively.

Variation

H1.2 deals with differences as to variation among the seniors, managers, and partners. Based on Figure 8, which plots the standardized materiality scores by level of responsibility, no conspicuous differences in variation about the apparent means among the seniors, managers, and partners are visible.

Similar to the analyses reported above with respect to location, effects associated with level of responsibility were examined within each firm. The sums of squares were calculated for the seniors, managers, and partners, separately for each firm; and Cochran's test for homogeneity of variance was applied to those sums of squares, on a
Summary of Level of Responsibility Effects

As described by the analyses above, one cannot reject the null hypothesis for either H1.1 (regarding location) or H1.2 (regarding variation). Overall, there were no real differences, either with respect to the means of the standardized materiality scores or with respect to the sums of squares associated with the standardized materiality scores, among the seniors, managers, and partners. These results are summarized in Table 10.

The one-way ANOVAs with respect to level of responsibility, calculated separately for each firm, identified only one significant difference. For firm #2, the seniors had higher materiality thresholds, on average, than the managers and partners (p ≤ .01). However, that was an isolated instance. Using Fisher's method for combining tests, the overall test statistic was significant at the .05 level (and was driven by the significant result with respect to firm #2), but it was not significant at the .01 level. Further, no effect associated with level of responsibility was identified as to location for either audit technology group, although the unstructured group could be viewed as having a borderline level of significance (p = .08).

Similar to the tests for differences as to location, the tests for differences as to variation were applied firm-by-firm. The seniors, managers, and partners were not different with respect to variation for any of the firms at the .05 level. The specific p-values inferred from Cochran's test for homogeneity of variance were
combined into an overall test statistic using Fisher's method. The overall test statistic was not significant at any reasonable level. Further, the seniors, managers, and partners were not significantly different with respect to variation for either of the two audit technology groups at the .05 level.

Based on these results, one may appropriately regard the seniors, managers, and partners as homogeneous, and proceed with the overall comparisons at the firm and audit technology levels. Those analyses follow.

**Firm Effects**

As described in Chapters II and III above, systematic differences among the participants' materiality judgments at the firm level were of interest. Prior research suggests that there may be differences between Big Eight and non-Big Eight firms or between national and non-national firms. Big Eight firms have been viewed as largely homogeneous due to the similarities in their client bases, national/international employee pool, sophisticated training, etc. Prior research has not explored whether there may be systematic differences in auditors' materiality thresholds associated with Big Eight firm affiliation. To address that issue, analyses similar to those described above with respect to level of responsibility were performed. Two hypotheses were explicitly stated with respect to firm
Figure 9

Standardized Scores for the Income Statement Materiality Judgments
Plotted by Firm Affiliation
Table 11
ANOVA Table For Firm Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Degrees of Freedom</th>
<th>Sums of Squares</th>
<th>Mean Square</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>5</td>
<td>12,599.5</td>
<td>2,519.9</td>
<td>10.07</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>66</td>
<td>16,512.4</td>
<td>250.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>71</td>
<td>29,111.9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Consequently, the existence of firm differences in location was fairly consistently observed across all three levels of responsibility. The overall firm differences reported above do not appear to be driven by the judgments of a single level.

Two separate one-way ANOVAs looked for firm differences as to location within each audit technology group. Among the three structured firms (firms #1-3), firm #1 was weakly significant (p=.043) relative to firm #3. (Since firm #1 had the highest mean and firm #3 had the lowest mean with respect to the structured firms, the p-value obtained in the one-way ANOVA can be attributed to the comparison of those two firms.) Among the three unstructured firms (firms #4-6), firm differences were not significant (p=.44). This result, in conjunction with the observation made earlier that the mean for each structured firm exceeded the mean for each unstructured firm, suggests that firm differences as to location were more conspicuous between firms having different audit technologies.

Variation

H2.2 deals with differences as to variation across the six participating Big Eight CPA firms. Based on Figure 9, which plots the standardized materiality scores by firm affiliation, differences in variation about the firms' means are visible. For example, the distribution for firm #1 is widely dispersed, whereas the distribution
This indicates that each of the structured firms was more variable than at least one of the unstructured firms.\textsuperscript{11}

Similar to the tests for location differences reported above, two types of supplemental comparisons with respect to variation were made for exploratory purposes. First, tests were conducted to identify firm differences in variation at each level of responsibility. Second, tests were conducted to identify whether there were any firm differences as to variation within each of the audit technology groups.

To test for differences in variation across firms, by level of responsibility, the variances of the standardized scores were calculated, by firm and by level of responsibility. (These variances are summarized in Table 12.) Cochran's test for homogeneity of variance was applied to these variances across firms, on a level-by-level basis.

For the seniors, significant differences in variation were detected among the firms. Cochran's test was applied sequentially, deleting each significant firm. Firm \#1 was the most variable, followed by firms \#2, 4, 3, 6, and 5 in order of decreasing variability. Firms \#1, 2, and 4 were significant at the .01 level, and firms \#3 and 6 were significant at the .05 level.

For the managers, significant differences in variation were also found among the firms. Firm \#1 was the most variable, followed by
firms #2 and 3. Each was significant at the .01 level; no other
differences among the managers were significant at the .05 level.

For the partners, significant differences in variation were also
identified among the firms. Firm #1 was the most variable, followed
by firms #3 and 6. Firms #1 and 3 were significant at the .01 level,
and firm #6 was significant at the .05 level. These were the only
differences among the partners that were significant at the .05 level.

The results of these analyses of firm differences, by level of
responsibility, are consistent with the overall firm effects reported
above. These analyses identified fairly pervasive firm effects across
each of the levels. The overall firm effects discussed above do not
appear to be driven by any particular level of responsibility.

Finally, two other tests looked for firm differences as to
variation within each audit technology group. Cochran's test for
homogeneity of variance was applied to the sums of squares for the
structured firms (firms #1-3 above) and, separately, to the sums of
squares for the unstructured firms (firms #4-6 above). No differences
in variation were detected for the three firms in the structured group
or for the three firms in the unstructured group at the .05 level.
This supplemental analysis suggests that the differences in variation
were much greater for firms having different audit technologies than
for firms having the same audit technology.
Table 13
Summary of Firm Effects

Comparisons Made | Re: Location | Re: Variation
---|---|---
A. Overall firm effect | Firm affiliation was highly significant (p ≤ .0001). Firm #1 was high relative to all other firms. | Firm differences were found at the .01 level. (Firms #1, 2, and 3 were high relative to firms #4–6.)

B. Firm affiliation, by level of responsibility | Firm affiliation was highly significant for the seniors (p ≤ .0001) and managers (p ≤ .001); the partners were borderline (p= .06). | Firm differences were found at the .01 level among the seniors, the managers, and the partners.

C. Firm affiliation, by audit technology | Firms #1 and 3 were weakly significant for the structured firms (p= .043); there were no differences among the unstructured firms (p= .44). | No difference was found at the .05 level for either audit technology.

Note: This table summarizes the analyses identified in Table 5. Table 5 summarizes the primary method of analysis for differences as to location and variation, identifies the data on which the above conclusions are based, and ties the analyses conducted to the notation introduced in Figure 7.
Firm differences were not significant for either of the two audit technologies. As was pointed out with respect to location, differences in variation at the firm level were more conspicuous for firms having different audit technologies than for firms within the same audit technology group.

These results indicate that there were real differences in the judgments of the participants at the firm level. In addition, the supplemental analyses conducted across firms within each audit technology suggested that differing audit technologies may have been a factor in the differences among the firms, both as to location and variation. The next section looks specifically at that issue.

**Audit Technology Effects**

In order to explore further the issue of systematic differences among the Big Eight firms, the participating firms were classified as structured or unstructured, based on the notion developed by Cushing and Loebbecke [1983] and Kinney [1986]. As discussed in Chapters II and III, this concept of audit technology has implications with respect to auditors and their differing materiality judgments. In particular, Cushing and Loebbecke commented on the divergent policies and practices with respect to planning stage materiality judgments across the firms that they studied. Kinney (in addition to Cushing and Loebbecke) pointed out that the structured firms provide more
unstructured firms). That plot is useful in identifying both differences among the six firms and differences among the two audit technology groups. Differences among the structured and unstructured firms are discussed first with respect to location, and second with respect to variation.

**Location**

H3.1 deals with differences as to location between the structured and unstructured audit technology groups. As mentioned, Figure 9 plots the standardized materiality scores by firm affiliation; there are visually apparent differences in the means for firms #1-3 (the structured firms) relative to those for firms #4-6 (the unstructured firms). Each of the structured firms has an apparently higher mean than each of the unstructured firms.

To test H3.1, a one-way ANOVA compared the standardized scores for systematic differences in the means for each audit technology group. Based on the 36 subjects from each audit technology group, the means of the standardized scores were 11.345 (-11.345) for the structured (unstructured) firms. As documented in Table 14, the ANOVA found audit technology to be highly significant (p \( \leq .0001 \)). On this basis, one can reject H3.1; there were real differences between the structured and unstructured firms with respect to the central tendencies of the materiality judgments.
That audit technology was highly significant is not surprising, in view of the results regarding firm effects reported in the preceding section. With respect to firm differences as to location discussed above, only firm #1 was shown to be significantly different than the other firms, based on a one-way ANOVA. However, the fact that the mean for each of the three structured firms exceeded the mean for each of the three unstructured firms strongly suggests that the overall significance of audit technology is not driven by the earlier reported significance of firm #1.

In order to explore the pervasiveness of the overall audit technology effect, an additional type of comparison was made. The purpose of this supplemental analysis was to ascertain whether the overall audit technology effect was consistently observed across the three levels of responsibility. Each level was examined separately to identify audit technology differences among the seniors, the managers, and the partners. Three separate one-way ANOVAs tested for audit technology differences between the structured and unstructured firms, on a level-by-level basis.

For the seniors, audit technology was highly significant ($p \leq 0.0001$). The seniors affiliated with the structured (unstructured) firms had a mean standardized score of 19.8 (-14.1). For the managers, audit technology was also highly significant ($p \leq 0.005$). The managers affiliated with the structured (unstructured) firms had a mean standardized score of 11.3 (-9.9). For the partners, audit
more variable than the participants from unstructured firms--95 percent of the total sums of squares was attributable to the structured firms (p \leq .01). Accordingly, one can reject H3.2, regarding equal variances between the two audit technology groups.\(^ {14} \)

In view of the analysis of firm effects reported in the preceding section, it is not surprising that the structured firms (#1-3) were found to be more variable than the unstructured firms (#4-6). With respect to firm differences in variation discussed above, Cochran's test was applied sequentially to the firms' sums of squares, deleting the most variable firm each time the firms were shown to have unequal variances. When all six firms were considered, firm #1 was the most variable and the firms were shown to have unequal variances. After deleting firm #1, firm #2 was the most variable and, again, the remaining firms were shown to have unequal variances. After further deleting firm #2, firm #3 was the most variable and, again, the remaining firms were still shown to have unequal variances. Yet, a supplemental analysis regarding variation was unable to demonstrate any statistically significant differences among the three structured firms and, separately, among the three unstructured firms.

In order to explore the pervasiveness of the overall audit technology effect with respect to variation, an additional type of comparison was made. The purpose of this supplemental analysis was to ascertain whether the greater variability of the structured firms was consistently observed across the three levels of responsibility. Each
level was examined separately to identify audit technology differences among the seniors, the managers, and the partners.

The sums of squares were calculated separately for the twelve subjects at each level of responsibility within each audit technology group. (These are summarized in Table 15.) Cochran's test for homogeneity of variance was used to test the equality of variance between the structured and unstructured firms, on a level-by-level basis. The seniors, the managers, and the partners affiliated with structured firms were, at each level, more variable than their counterparts affiliated with unstructured firms (p \leq 0.01). The overall greater variability of the structured firms was consistently observed at each level of responsibility.

Summary of Audit Technology Effects

Audit technology effects were identified with respect to both location and variation. Accordingly, both H3.1 (regarding differences as to location) and H3.2 (regarding differences as to variation) were rejected. These results are summarized in Table 16.

With respect to location differences, audit technology was highly significant (p \leq 0.0001), based on the 36 subjects representing each of the two audit technology groups. Participants from the structured firms had higher materiality thresholds relative to the participants from the unstructured firms. Indeed, the mean associated with each of the structured firms was higher than the mean associated with each of
Table 15

Sums of Squares of the Standardized Scores
By Level of Responsibility and By Technology

<table>
<thead>
<tr>
<th>Level of Responsibility</th>
<th>Structured Firms</th>
<th>Unstructured Firms</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Senior</td>
<td>3424</td>
<td>217</td>
<td>3641</td>
</tr>
<tr>
<td>Manager</td>
<td>5529</td>
<td>211</td>
<td>5740</td>
</tr>
<tr>
<td>Partner</td>
<td>8164</td>
<td>432</td>
<td>8596</td>
</tr>
<tr>
<td>Total</td>
<td>17,117</td>
<td>860</td>
<td>17,977</td>
</tr>
</tbody>
</table>
Table 16
Summary of Audit Technology Effects

<table>
<thead>
<tr>
<th>Comparisons Made</th>
<th>Re: Location</th>
<th>Re: Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Overall audit technology effect</td>
<td>Audit technology was highly significant (p ≤.0001). The structured firms had higher thresholds than the unstructured firms.</td>
<td>The structured firms had greater variability than the unstructured firms (p ≤.01).</td>
</tr>
</tbody>
</table>

B. Audit technology, by level of responsibility
Audit technology was highly significant for the seniors (p ≤.0001) and for the managers (p ≤.005); the partners were not significant (p=.12). At each level, the structured firms had higher thresholds.

Note: This table summarizes the analyses identified in Table 6. Table 6 summarizes the primary method of analysis for differences as to location and variation, identifies the data on which the above conclusions are based, and ties the analyses conducted to the notation introduced in Figure 7.
the materiality judgment, H4.2 dealt with the association between the auditors' risk attitudes and the extremeness of their materiality thresholds.\^{15} It was expected that auditors who were relatively more concerned with Type I errors (e.g., practice development) would tend to have higher materiality thresholds, whereas auditors who were relatively more concerned with Type II errors (e.g., litigation by financial statement users) would tend to have lower thresholds.

As explained in Chapter III, risk attitude was based on an expectancy-valence measure with respect to the four cells identified in Figure 3. (The individual items associated with the four cells were identified in Table 7.) Subjects were asked to specify their beliefs (likelihoods on a seven-point scale) and their evaluations (assessments of importance on a five-point scale). The product of the belief and evaluation was calculated for each item (termed the item's expectancy-valence), for each subject. The items' expectancy-valences were summed within each cell for each subject.

Risk attitude was defined to be the subject's tradeoff of Type I and Type II errors. It was measured by the expectancy-valence for cell #4 (regarding Type II errors) divided by the expectancy-valence for cell #1 (regarding Type I errors) for each subject.

To construct a reasonably powerful test to identify any differences among the participants as to risk attitude, and to measure the strength of the association between risk attitude and their materiality judgments, subjects were partitioned into one of three
materiality threshold group would emphasize cell #1 regarding Type I errors.) On the basis of the ANOVAs with respect to the overall risk attitude measures and with respect to the risk attitude measures calculated separately for Type I and Type II errors, H4.1 (that there are no systematic differences among the auditors with respect to the risk attitude measures) cannot be rejected.\textsuperscript{16}

Using the Pearson product-moment correlation, subjects' risk attitude measures were correlated with their standardized materiality scores. The overall correlation coefficient between the risk attitude measures and the standardized materiality scores was .21203, and was not significant at the .05 level (p=.074). It was surprising that this coefficient was positive, since a negative association was hypothesized. (The p-value of .074 was comparable in magnitude to the overall significance reported by Kida [1980] between his expectancy-valence measure of attitude and his subjects' qualifying behavior--p = .068. However, as indicated, the observed direction was opposite to that hypothesized here.)\textsuperscript{17}

These results suggest that the null hypothesis for H4.2 cannot be rejected (which is not surprising since the null hypothesis for H4.1 could not be rejected either). The measure of risk attitude adopted in this study was not found to vary systematically among the subjects, nor was it systematically related to participants' materiality thresholds. The overall notion of risk attitude, as operationalized
Table 17

Attitudinal Differences (p < .10) Among Subjects With the Lowest Versus Highest Thresholds

<table>
<thead>
<tr>
<th>Cell of Fig. 3.1</th>
<th>Item</th>
<th>Belief</th>
<th>Eval'n</th>
<th>Exp.Val.</th>
<th>Highest</th>
<th>Lowest v.</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Lose the client</td>
<td>.0555</td>
<td>.0937</td>
<td>L&gt;H</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Litigation by client</td>
<td></td>
<td></td>
<td>.0192</td>
<td>L&gt;H</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Damage own standing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Enhance firm's reputation</td>
<td>.0203</td>
<td>.0319</td>
<td>L&gt;H</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Damage client's reputation</td>
<td>.0797</td>
<td>.0210</td>
<td>L&gt;H</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Deteriorated client relations</td>
<td>.0873</td>
<td>.0494</td>
<td>L&lt;H</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Damage client's reputation</td>
<td>.0212</td>
<td></td>
<td>L&lt;H</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
finally, several comparisons were made to identify whether the income statement materiality thresholds were less than the balance sheet materiality thresholds, as expected on average (for each level of responsibility, for each firm, and for both audit technologies). These comparisons were made, on a case-by-case basis, to explore ex post the appropriateness of viewing the income statement materiality judgments as the dependent variable of primary interest.

**Individual ANOVA Models**

In order to gain insights into the nature of individual differences among the participants, ANOVA models were calculated for each subject across the 24 cases.18 These models were based on subjects' untransformed (or "raw") judgments; and they were intended to be descriptive, not to represent the actual "processes" which subjects followed in arriving at their judgments.

In the aggregate, for all 72 subjects, income (loss) before taxes ("income") was significant for 49 subjects. The total asset variable was the second most frequently significant variable, and was significant for 24 subjects. The sales variable was significant for 15 subjects. Only one interaction was often significant; income x assets was significant for 19 subjects. The seniors tended to rely more on sales relative to the managers and partners, who apparently relied on income to a greater extent. Table 18 summarizes the
significant cues (p ≤ .05), by level of responsibility and, separately, by firm affiliation.

Table 19 summarizes the significant cues for the subjects for whom only one variable was significant. Income was the only significant variable for 29 subjects, followed by sales for eleven subjects, and total assets for seven subjects.

Firm differences were identifiable. Total assets was significant for 24 subjects. Nearly half (eleven) were from firm #1, a structured firm. In addition, all seven of the subjects for whom total assets was the only significant variable were affiliated with firm #1. Further, only one subject from firm #1 for whom just one variable was significant relied on a variable other than total assets (i.e., income).

The standardized materiality scores for the seven subjects from firm #1 for whom total assets was the only significant variable were examined further. Of the seven highest standardized scores, six were subjects from firm #1, and all six had only one significant variable—total assets. These subjects comprised half of the twelve subjects from firm #1. Apparently, the conspicuously high materiality thresholds associated with subjects from firm #1 were driven, or at least influenced, by their relatively strong balance sheet orientation within the context of this experiment.

Although somewhat ad hoc, these observations underscore the existence of individual differences, even in the structural form of
the models of subjects' judgments. This is somewhat surprising in view of the fairly limited experimental manipulations—only three financial statement cues were available and only two of these were directly applicable to the income statement. Particularly with respect to firm #1, differences in subjects' thresholds apparently related to differences in subjects' structural models.

Business Risk Assessments

To the extent that subjects perceived the financial statement cues as implying different degrees of business risk (i.e., exposure) to the auditors, their materiality thresholds may be affected. All else equal, higher exposure to the auditor is believed to result in a lower materiality threshold. To ascertain whether there was such a confounding effect present in the experiment, subjects were asked to rate the business risk associated with each case on a seven-point scale (where "1" represented "very low" and "7" represented "very high"). The responses were analyzed by one-way ANOVAs, on a case-by-case basis, to evaluate whether there were systematic differences which might affect the above-reported results with respect to level of responsibility, firm affiliation, or audit technology.

With respect to level of responsibility, only two cases were significant at the .05 level. The seniors expressed a higher level of risk associated with case #5 than did the managers (p=.048). The
The responses for the repeated cases were compared to the original responses in three ways. First, the overall Pearson product-moment correlation coefficient was calculated to identify whether the responses on the four repetitions moved in the same direction as those to the original cases. Second, paired t-tests were calculated, on a case-by-case basis, to identify whether, on average, the responses to the four repeated cases were comparable to the responses to the original cases. Last, separate one-way ANOVAs were run on the differences between the responses to the original and repeated cases to identify any systematic differences among the levels of responsibility, among the firms, and between the two audit technologies.

Across all 72 subjects, the correlation coefficient of the repeated cases relative to the original cases was approximately .963. Not surprisingly, larger responses on the original cases appear to be paired with larger responses on the repeated cases.

For the 72 subjects in the aggregate, the hypothesis of no difference between the repeated and the original cases could not be rejected for any of the four judgments. The resulting approximate t-values (p-values) for cases #3, 9, 18, and 22 were: -.36 (.72), -.96 (.34), -.67 (.50), and .73 (.47), respectively.

Separate one-way ANOVAs, with respect to level of responsibility, firm affiliation, and audit technology, were run on the differences between the original and repeated cases, on a case-by-case basis.
With respect to location, a one-way ANOVA of the standardized scores did not find any difference in the mean judgments between the two orders \((p=.71)\). With respect to variation, Cochran's test for homogeneity of variance (using the sums of squares associated with each order) did not detect a significant difference at the .05 level. Apparently, the judgments were not influenced by the order in which the cases were presented.

Income Statement versus Balance Sheet Judgments

SAS No. 47 (paragraph 12) describes the preliminary judgment about materiality, at the financial statement level, in terms of the smallest amount material to any one of the financial statements. Prior research has consistently reported net income (or income before taxes) to be the dominant variable in numerous studies of auditors' evaluation stage materiality judgments. Balance sheet items (such as total assets) have tended to be unimportant variables in those studies. Further, others have asserted that the planning stage judgments are generally driven more by income statement considerations rather than balance sheet considerations (e.g., Zuber, et al. [1983, p. 43]). Consequently, it was expected that the income statement materiality judgments would generally be less than the balance sheet materiality judgments (i.e., the preliminary judgment about materiality would result from income statement considerations) for most subjects for most cases. That expectation was examined in a
sheet materiality judgments for 23 of the cases for the structured firms and for all 24 of the cases for the unstructured firms.

These observations provide further (albeit ex post) justification that, for purposes of examining level of responsibility, firm, and audit technology effects, the income statement materiality judgments were appropriately viewed as the dependent variable of principal interest. In the preponderance of cases, the income statement materiality judgments, on average (across the three levels of responsibility, across the six firms, and between the two audit technology groups), would indeed form the basis for the auditors' "preliminary judgments about materiality."

Comments on Statistical Assumptions

The analyses reported above dealt with systematic differences among the auditors with respect to location (i.e., central tendency) and variation (i.e., dispersion) in their materiality judgments. Analysis of variance was the primary statistical method used to identify differences as to location. Cochran's test for homogeneity of variance was the primary method used to identify differences as to variation. Both are parametric procedures involving certain distributional assumptions about the data.

As discussed in Chapter III, ANOVA assumes that the data are normally distributed with independent error components having
The primary significant findings with respect to variation were corroborated in two ways. Differences at the firm and audit technology levels were also identified (as reported in footnote #9) using the conservative Bonferonni simultaneous confidence intervals, involving all possible comparisons (i.e., ratios) of the firms' sums of squares. Since these ratios constituted F-tests, the above remarks with respect to the robustness of the F-test apply with respect to the tests of variation, as well. Again, distinct firm and audit technology effects were observed. Finally, nonparametric Wilcoxon rank sum tests were used to further examine the differences in variation at the firm and audit technology levels (reported in footnotes #8 and 12, respectively). Similar results obtained.

Based on the robustness of the methods used (i.e., the F-tests used in the ANOVAs with respect to location and used to corroborate the results with respect to variation) and based on the substantiation provided by the nonparametric tests, no further analyses related to the underlying statistical assumptions are warranted. The results were consistently observed, and apparently do not hinge on the distributional assumptions underlying the parametric methods that were used in the data analysis.
task. The seniors, managers, and partners were identified by the codes "2," "3," and "4," respectively, related to the basic hierarchy of the firms.

A few subjects were visibly more extreme than others in their materiality judgments, as demonstrated by Figure 8. However, the seniors, managers, and partners overall were not apparently affected differentially due to these individuals. (Further, it should be noted that there are 24 subjects at each level of responsibility, and individual differences that are not too extreme will be mitigated when averaged across the other subjects within each level of responsibility.) None of the observations are conspicuously extreme among the seniors, managers, and partners. Accordingly, the results of the analyses are probably not greatly influenced by "outliers."

5. Similar results were found using (nonparametric) Wilcoxon rank sum tests. Within firm #2, the judgments of the seniors were significantly higher than those of both the managers and partners ($p = .014$ in each instance). These results suggest that the (parametric) ANOVA results are not heavily influenced by the presence of a few relatively extreme observations.

Table 8 indicates that the seniors from firm #2 had fewer years of experience, on average, than those from each of the other firms. However, based on pooled t-tests, there were no significant differences in years of experience for the seniors across the six firms at the .01 level. (Since there were five comparisons involving firm #2, the .01 level for each comparison corresponds to an overall .05 significance level.) Conceivably, judgments of the seniors from firm #2 may be influenced by a relative lack of experience in making planning stage materiality judgments. To the extent that the experimental task involved subjects' "rules of thumb" with respect to materiality judgments at the aggregate financial statement level, these subjects' heuristics may be less well defined than other subjects having more practice with the real world analogue of the experimental task.

6. A two-way ANOVA encompassing the main effects for level of responsibility and firm affiliation (as well as the level-firm interaction) was conducted. The F-values (p-values) for level of responsibility, firm affiliation, and the level-firm interaction were as follows: 1.22 (.3041), 11.92 (.0001), and 2.17 (.0340).

At first glance, the significant interaction might suggest that firm effects cannot be discussed meaningfully without considering specific levels of responsibility. However, the significant interaction resulted entirely from the previously reported significant effect (regarding the seniors) in firm #2.
Nonparametric Wilcoxon rank sum tests were applied to these variances. Specifically, all pairwise comparisons were made across the six firms using Wilcoxon rank sums (there were three observations per firm, representing the variances of the seniors, managers, and partners for each firm). Firm #1 was significantly different than each of the other firms (p < .05); none of the other firms were significant at the .05 level. This provides further evidence of differences at the firm level in variation, and corroborates the parametric results reported in the text. An important aspect of this nonparametric result is that no distributional assumptions are required.

11. Further evidence of firm differences as to variation resulted from all pairwise comparisons using the ratios of the firms' sums of squares. A total of fifteen comparisons resulted; three were comparisons between structured firms, three were comparisons between unstructured firms, and nine were comparisons involving one structured and one unstructured firm. Significance was based on the conservative Bonferonni simultaneous confidence intervals involving the .05 level.

None of the comparisons among the structured firms and none of the comparisons among the unstructured firms were significant at the .05 level. However, seven of the nine pairwise comparisons between a structured and an unstructured firm were significant (and another could be viewed as borderline). This provides strong evidence that firm differences in variation were closely related to their audit technologies. This issue will be discussed in more depth in a subsequent section, entitled "Audit Technology Effects."

12. Chapters II and III discussed the Cushing and Loebbecke [1983] and Kinney [1986] papers with respect to audit technology. As noted in Chapter III, Kinney classified Deloitte, Haskins, & Sells; Peat, Marwick, Mitchell; and Touche Ross as the "structured" firms, and they are so classified here. Further, Kinney classified Coopers & Lybrand and Price Waterhouse as "unstructured" firms, and they are so classified here. In addition, Kinney classified Arthur Andersen, Arthur Young, and Ernst & Whinney as "intermediate" firms. As explained in Chapter III, the latter two firms were not involved in this study, and Arthur Andersen was included among the unstructured firms. (Based on a review of the firms' audit manuals and discussions with each "contact person," Arthur Andersen's policies with respect to planning stage materiality judgments appeared to be similar to those of Coopers & Lybrand and Price Waterhouse, at the time the experiment was conducted.)

This dissertation relies on Kinney's classification to compare the most highly structured firms with others viewed as less structured. The classification of Arthur Andersen among the "unstructured" firms is somewhat subjective, and the cut-off between
and audit technology; no systematic differences in the risk attitude measures were identified at the .05 level.

17. Similar correlations were calculated by level of responsibility and by firm affiliation. By level of responsibility, the correlation coefficients (p-values) for the seniors, managers, and partners were, respectively: -.08440 (.6950), .40335 (.0506), and .28073 (.1839). By firm, the coefficients (p-values) for firm #1-6 were, respectively: .21125 (.5099), .08309 (.7974), .47506 (.1186), .19222 (.5495), -.08212 (.7997), .17512 (.5862). None of the coefficients, by level of responsibility or by firm, were significant at the .05 level. In addition, nearly all of the coefficients were opposite the direction hypothesized.

18. This section discusses the individual ANOVA models with respect to the income statement materiality judgments. Individual ANOVA models were also constructed with respect to the balance sheet materiality judgments. Those are discussed in Appendix D.

19. The firm's guidelines emphasize owners' equity, a balance sheet concept, in the determination of materiality at the financial statement level. Since the case materials provided the breakdown of the balance sheet in percentage terms, owners' equity could be ascertained from the total assets variable.

20. The possibility must be acknowledged that the extremeness of these standardized scores may be an artifact of the experiment. To the extent that firm #1, a structured firm, bases the planning stage materiality judgment on owners' equity (a balance sheet amount), it may be particularly abstract for subjects of firm #1 to express separate judgments with respect to the income statement and balance sheet. In addition, this firm may be affected more than the others by the orthogonal experimental design. In the experiment, total assets, income, and sales were uncorrelated as a consequence of the full factorial design. However, the "contact person" (a partner) stated that his colleagues would be able to provide meaningful responses based on his review of the experimental materials. In addition, none of the subjects from firm #1 expressed difficulty with the experimental task in the survey questionnaire.

21. This section discusses the analyses with respect to the repeated cases for the income statement materiality judgments. Similar analyses were performed with respect to the repeated cases for the balance sheet materiality judgments. Those are discussed in Appendix D.

22. The survey questionnaire (Appendix C) contained a few questions which related to the level of interest subjects had toward
represented by the y-axis and the case number is represented by the x-axis).

The plot of the calculated ratios did not indicate any tendency for the subjects to become systematically stricter (or more lenient) as they completed the experiment. Further, Kendall's nonparametric test for independence (Hollander and Wolfe [1973, pp. 185-6]) indicated that the ratios were not significantly associated with the order of case presentation at the .05 level.
SAS No. 47 indicates that the auditor should consider the preliminary judgment about materiality at the planning stage of an audit engagement. Further, the SAS suggests that the judgment may be made at the aggregate financial statement level; the preliminary judgment about materiality would then be the smallest amount material to any one of the financial statements. On that basis, subjects were asked to specify the appropriate planning stage materiality threshold at the aggregate financial statement level, separately with respect to the income statement and balance sheet, for each of 24 different cases. The 24 cases represented a full factorial design related to three financial statement variables—-income (loss) before taxes (with four levels), net sales (with two levels), and total assets (with three levels).

Based on prior research, primarily involving auditors' evaluation stage materiality judgments, and based on the professional literature with respect to auditors' planning stage materiality judgments, the income statement materiality judgments constituted the dependent variable of principal interest. Accordingly, the results discussed in Chapter IV dealt with these income statement materiality judgments. (Similar analyses were conducted with respect to the balance sheet materiality judgments for exploratory purposes, and those results were presented in Appendix D for the interested reader.)

In order to mitigate the conspicuous heterogeneity of variance across the individual cases, and to provide a single, independent
statistic, based on Fisher's method, was significant at the .05 level (but was not significant at the .01 level). Across the firms, differences among levels of responsibility were not pervasive. The results did not suggest any real differences among the three levels of responsibility, and, accordingly, H1.1 was not rejected.

With respect to variation, no significant differences were detected among levels of responsibility for any of the firms, using Cochran's test for homogeneity of variance. Not surprisingly, therefore, the overall test statistic, based on Fisher's method, was not significant at any reasonable level. Accordingly, H1.2 was not rejected, either. Consequently, the seniors, managers, and partners were viewed as homogeneous for purposes of investigating firm effects and, related to that, audit technology effects.

Two hypotheses were stated with respect to overall firm effects. H2.1 dealt with differences among the six firms as to location, and H2.2 dealt with differences as to variation.

With respect to location, overall firm differences were conspicuous. Based on a one-way ANOVA, firm affiliation was highly significant (F=10.07, p <.0001). In particular, firm #1 was identified as having a significantly higher mean materiality threshold than each of the other firms. On that basis, the null hypothesis for H2.1 (that there were no differences in location among the six participating firms) was rejected. Interestingly, each of the
With respect to location, a difference in the mean thresholds between the two audit technology groups was found. Based on a one-way ANOVA, audit technology was highly significant ($F=32.69$, $p < .0001$). On that basis, the null hypothesis for H3.1 (that there were no differences in location between the structured and unstructured firms) was rejected. That the structured firms had significantly higher materiality thresholds was not surprising, since the analysis of firm effects indicated that each of the structured firms had higher mean standardized materiality scores (although the differences were not necessarily significant) relative to each of the unstructured firms. In view of the relatively uncomplicated judgment context (i.e., relative to the "real world"), this result is particularly striking.

With respect to variation, a difference between the two audit technology groups was also identified. Cochran's test for homogeneity of variance was applied to the sums of squares associated with each of the audit technology groups. The participants affiliated with structured firms were significantly more variable in their judgments than the participants affiliated with unstructured firms ($p < .01$).

Further compelling evidence of audit technology differences came from all pairwise comparisons using ratios (i.e., F-tests) of the firms' sums of squares. Of the fifteen possible pairwise comparisons, none of the six comparisons within the same audit technology group was significant. However, seven of the nine comparisons between firms...
financial statement users) were expected to have relatively low materiality thresholds.

To construct a fairly powerful test to identify differences among subjects as to their risk attitudes and to measure the association with their materiality judgments, subjects were partitioned into thirds, based on the extremeness of their standardized scores. The subjects in the upper and lower thirds were compared for systematic differences in risk attitudes, and their risk attitudes were correlated with their standardized scores. However, the subjects did not differ systematically with respect to the measure of risk attitude adopted. (Further, across all 72 subjects, risk attitudes did not differ when comparisons were made in terms of level of responsibility, firm affiliation, and audit technology.) Hence, H4.1 was not rejected.

Similarly, the measure of risk attitude adopted in the study was not significantly associated with auditors' materiality thresholds at the .05 level. Hence, H4.2 was not rejected. As a result, the measure of risk attitude adopted was not helpful in explaining systematic differences among the subjects with respect to their planning stage materiality judgments.

To further investigate the issue of individual differences with respect to the materiality judgments, an ANOVA model was constructed for each subject across the 24 cases. Even in the structural forms of the ANOVA models (encompassing each main effect and each two-way
respect to level of responsibility, firm affiliation, or audit technology.

Two other analyses were conducted to assess the meaningfulness of subjects' responses. The first involved test-retest consistency and the second checked for order effects. No significant differences were detected, in the aggregate, between subjects' responses to the original cases and those to the four repeated cases. Further, no systematic differences were identified with respect to level of responsibility, firm affiliation, or audit technology. Similarly, no differences were detected between the two orders of case presentation, either in terms of location or variation. Apparently, the subjects were able to make meaningful responses; at least, no evidence was obtained to the contrary.

Finally, comparisons were made, on a case-by-case basis using subjects' untransformed responses, between the income statement and balance sheet materiality judgments to explore the extent to which, on average (in terms of level of responsibility, firm affiliation, and audit technology), the auditors' preliminary judgments about materiality were driven by income statement considerations in the context of the experiment. For the preponderance of the cases, the average income statement judgment for each level of responsibility, for each firm, and for both audit technology groups, was lower than the corresponding balance sheet judgments. These observations were consistent with the expectation that the income statement would be the
may not be surprising, since SAS No. 47 implies that the planning stage judgments depend primarily on quantitative considerations, while the evaluation stage judgments are influenced to a greater extent by qualitative considerations. Others (e.g., Woolsey [1954a] and Pattillo [1976]) have suggested that, in an evaluation stage context, there may be a range such that an item below some threshold is clearly immaterial, an item above some other threshold is clearly material, and an item between those two thresholds must be evaluated with respect to its qualitative characteristics in order to assess materiality. The experimental task may have dealt with such quantitative thresholds, and subjects' responses may have been determined by their heuristics adopted to assist them in their tentative materiality judgments. Perhaps level of responsibility (presumably related to experience) plays a greater role in identifying and incorporating the various qualitative considerations in an evaluation stage materiality judgment when the quantitative considerations are ambiguous.

Substantial differences were identified at the firm level among the six participating Big Eight firms, both as to location and variation. Prior studies of auditors' (evaluation stage) materiality judgments have pointed out differences between national and non-national (regional) firms or, alternatively, between Big Eight and non-Big Eight firms. The Big Eight firms have been viewed as so similar as to be regarded as homogeneous organizations. This study
location and variation. The structured firms had higher thresholds and greater variability relative to the unstructured firms.

That the structured firms had higher materiality thresholds was not surprising. By imposing firm-wide guidelines throughout the organization, the structured firms were expected to manage the risks associated with Type II errors (i.e., a portfolio approach to risk management), while allowing them to have relatively higher materiality thresholds. However, to employ that explanation of their higher materiality thresholds, the structured firms must have lower variability at the same time.

Therefore, the greater variability of the structured firms relative to the unstructured firms was surprising, and one can only speculate about why the structured firms were more variable in their judgments. Perhaps the nature of the judgment task was not amenable to the structured firms' guidelines. The experimental task required subjects to make their materiality judgments at the aggregate financial statement level, separately for the income statement and balance sheet. The structured firms' guidelines do not make the distinction between the income statement and balance sheet judgments. However, the contact people at the structured firms were explicit that such a distinction would not be problematic for participants from their firms.

Alternatively, the participants from the structured firms may have departed from their firm guidelines and responded as they
departed from those guidelines), that subjects from structured firms may become even more variable in their judgments than other auditors from firms not having such detailed guidelines in the first place is an interesting consideration. This issue has important practical consequences to the structured firms, since such a result is obviously inconsistent with the ostensible purpose of imposing a more structured approach within a firm.

The notion of differing risk attitudes (as an intervening variable) was not helpful in understanding why subjects differed in their materiality thresholds. The overall risk attitude measures were not significantly correlated with subjects' standardized materiality scores. Similarly, the separate measures related to Type I and Type II errors were not significantly associated with the materiality judgments. Nevertheless, there were a few significant differences (regarding specific items incorporated into the aggregate risk attitude measure) identified between the subjects having the highest and lowest materiality thresholds, but the differences were not clear-cut.

One can only speculate about why there was a tendency for subjects' risk attitude measures, related to Type I and Type II errors, to be in directions opposite to that hypothesized. (Subjects with relatively low materiality thresholds were expected to have higher measures related to Type II errors; instead, the subjects with high materiality thresholds had higher expectancy-valences related to
Limitations of the Experiment

Experience was operationalized as level of responsibility, i.e., senior, manager, or partner. The null hypotheses related to HI.1 (regarding differences as to location) and HI.2 (regarding differences as to variation) could not be rejected. Overall, level of responsibility was apparently unrelated to differences among the subjects' materiality judgments. Since experience could be operationalized in other ways (for example, based on years as an auditor, or based on the number of times actual planning stage materiality judgments have been quantified), there can be no assurance that other experience-related constructs would lead to the same results.

Firm affiliation involved the (six) individual Big Eight firms participating in the study. The tests of audit technology effects compared the three firms viewed as the most structured (the "structured" firms) relative to the three firms viewed as less structured (the "unstructured" firms). The classifications with respect to structure relied substantially on Kinney [1986]. However, the classification of Arthur Andersen among the "unstructured" firms was somewhat subjective, since the cut-off between the "structured" and "unstructured" firms along the continuum of firms' degrees of structure was, to some extent, arbitrary.

Moreover, the results cannot be extrapolated to the two Big Eight firms who did not participate, and the effect on the results of
Similar to other behavioral experiments, this study may be criticized because it is not necessarily generalizable to the "real world" of auditing. Not all the variables potentially relevant to an auditor's planning stage materiality judgment could be incorporated, and no claim is made that the selected independent variables are comprehensive. In this study, the auditors were limited to three financial statement cues (although percentage information was provided about the composition of the income statement and balance sheet). In addition, an auditor would have much more detailed knowledge about an actual client than can be effectively conveyed about a hypothetical client.

Moreover, since a fixed-effects ANOVA model was used instead of a random-effects model, there can be no generalization beyond the specific levels chosen for the independent variables. Further, external validity may be sacrificed by employing an orthogonal experimental design. In real auditing situations, auditors deal with financial statement variables that are correlated; in the experimental materials, the cues were independent. Cues which have realistic magnitudes individually may not necessarily be realistic when combined across all levels of the other variables. However, subjects' comments on the survey questionnaire suggested that most subjects viewed the experimental materials as reasonably realistic. Only one subject, a partner with an unstructured firm, remarked that certain ratios were rather extreme for some combinations of the financial statement
There were a number of assumptions associated with analysis of variance and Cochran's test for homogeneity of variance, which were heavily used in this study. With respect to ANOVA, the assumptions focused on the residuals of ANOVA models. The residuals were assumed to be independent across observations, and to be distributed normally with mean zero and constant variance. However, these assumptions were not viewed as critical to the interpretation of the results. The strong corroboration from the nonparametric analyses suggested that the results were not dependent on the distributional assumptions of the parametric analyses. Similarly, Cochran's test primarily assumed that the data were normally distributed. The strong corroboration obtained from F-tests (robust with respect to non-normality) involving all pairwise comparisons among firms' sums of squares and similar results based on nonparametric Wilcoxon rank sums, strongly suggested that the results were not dependent on the distributional assumptions underlying the parametric methods.

Finally, whether (or how) differences among auditors in their preliminary judgments about materiality will affect their performance of an audit (or their "final" materiality judgments) is unclear. Conceivably, differences among auditors in their materiality judgments at the planning stage may be reduced or eliminated at the evaluation stage, following fieldwork. Therefore, such differences may or may not imply "quality" differences. This issue is among numerous areas for future research.
research should include participants from other national firms and, perhaps, from non-national firms. Such an extension would provide insights into the influence of differing audit technologies (i.e., structured versus unstructured approaches) across a wider population of CPA firms than just the Big Eight firms.

Additional investigation is warranted with respect to understanding why auditors establish different materiality thresholds. As reported above, there may have been some connection in this study between the cues utilized and the extremeness of the thresholds. However, the notion of differing risk attitudes, as operationalized here, was not helpful in clarifying why auditors' judgments differed. This is still an area about which little is known.

Subsequent research should include additional quantitative and qualitative variables as factors influencing auditors' planning stage materiality judgments. As mentioned, this could be accomplished in conjunction with limiting the number of cases to one or a very few. For example, a potentially interesting variable involves the risk characteristics of the client organization. This could encompass the distinction between public and private companies, since there is evidence that public companies are perceived as yielding greater exposure to auditors, all else equal.

An important topic not addressed in this dissertation concerns the impact which different planning stage materiality judgments might have on the conduct of an audit. Prior research has not explored what
APPENDIX A

INSTRUCTIONS AND BACKGROUND INFORMATION

Introduction

This experiment deals with the quantification of planning stage materiality judgments (termed "preliminary judgments about materiality") as described by SAS No. 47, "Audit Risk and Materiality in the Conduct of an Audit." As discussed in the SAS, materiality judgments arise in planning the audit engagement and in evaluating the accumulated audit evidence. This experiment focuses on the materiality judgment at the planning stage. Although prior "laboratory" experiments have reported differences in auditors' materiality judgments, no attempt has been made to understand why, within the context of those experiments, the judgments were different. Therefore, this experiment and the related questionnaires address some reasons why auditors may specify different materiality thresholds in such experiments.

In addition to planning stage materiality judgments, this experiment asks you to estimate the required audit hours (by staff, senior, manager, and partner) based on the information provided. Further, you are asked to express your perception of "business risk" (on a seven-point scale where "1" represents very low and "7" represents very high) for each of the individual cases. You may
is not provided; and the information that is provided lacks much of
the detail that would be available with respect to an on-going audit
client.) However, in planning actual engagements, auditors may
sometimes find that some desired information is unavailable.
Moreover, in this experiment, the relative levels of your judgments
are of more interest than the absolute levels. Consequently, these
materials invoke an assumption of "all else constant" to the extent
some information has been excluded. It is understood that judgments,
such as those you are being asked to make, are almost always tentative
and are revised as additional relevant information becomes available.

Even though the information presented for the hypothetical client
has been modeled from real companies, there are no "right" answers,
since the study deals with judgments (e.g., about materiality).
Consequently, your responses will not be evaluated for "correctness."
Nevertheless, if the study is to be valid, your responses must be
conscientious, and it is essential to complete each of the
requirements presented in the second booklet. Your responses will be
held in the strictest confidence; neither your identify, nor your
firm's, will be divulged in the discussions of the experiment's
results. Your assistance is greatly appreciated.

Instructions

SAS No. 47 does not require auditors to quantify their
preliminary judgments about materiality (i.e., planning stage
about materiality for the balance sheet and a corresponding single value for the income statement. If you are accustomed to providing a range of values for these judgments, please specify the range which you believe to be applicable; in addition, please specify the one value for the balance sheet and corresponding one value for the income statement which you deem most appropriate.

In addition, you are asked to complete a questionnaire that deals with auditors' attitudes toward qualified and unqualified opinions and their perceptions of the consequences associated with those opinions. Finally, you are asked to respond to a "debriefing" questionnaire that asks for some additional information which is commonly requested in experimental studies such as this. It is very important that you respond independently of others who may be participating; please do not discuss the experiment with others, in order to maintain the integrity of individuals' responses.

Background Information

Your role

You are the partner in charge of your firm's audit of the financial statements of Amicus Incorporated as of January 31, 1986, and for the year then ended. As you approach the planning of the engagement, a question has been raised about the appropriate planning level materiality threshold (for both the balance sheet and the income statement) which will be specified in the planning memorandum.
nearby communities. The relative sales volume for the stores is: downtown (48%), north-end (36%), branch #1 (10%), and branch #2 (6%).

Competition

Competition is strong in the retail industry and there are few barriers to entry. The major competitors are other department stores, some of which are large national companies. Additional pressure comes from numerous specialty stores, which are increasing in importance in the retail industry. Nevertheless, Amicus remains well-positioned, due in part to strong customer identification. Good planning by management led to the development of strength in several areas associated with the changing demographics and tastes of its customers.

Audit history

Your firm has audited the financial statements of Amicus since the fiscal year ended 1/31/79. That first audit was a consequence of a bank's requirement for an annual audit when Amicus borrowed $6 million for ten years in March of 1979. Prior to that first audit, your firm had provided tax services to Amicus for nearly 15 years.

Your firm has given an unqualified opinion each year, except for a qualified opinion the first year due to inability to verify the beginning inventory. Management has been willing to record your firm's proposed adjustments; in your view, the net effect of these adjustments has been minor in the past. Adjustments have usually
As in prior years, the two most significant areas of audit concern are inventory and receivables. Inventory is about 41 percent of total assets and is accounted for by the LIFO retail method. The client plans to close the four stores for two days (February 1 and 2) in order to conduct the annual physical inventory. You plan to verify, on a test basis, the client's physical counts, unit costs, and extensions.

Amicus has its own charge card, and trade receivables are its own charge accounts. About 21 percent of total assets at year-end consist of these receivables. Major planned audit procedures include confirmation of selected accounts and review of cash receipts subsequent to year-end.

Fees

You have compared the audit and tax fees paid by Amicus with those paid by other clients of comparable size and audit effort, and have found that the Amicus fees are about average, as a percentage of standard (desired) fees, relative to those of the other clients. Due to management's moderate resistance to fee increases, it is expected that this condition will continue in future years.

Composition of financial statements

As explained earlier, the cases presented in the second booklet differ as to the magnitudes of total assets, net sales, and income
APPENDIX B
ATTITUDE MEASUREMENT QUESTIONNAIRE

The purpose of this questionnaire is to elicit your attitudes toward qualified audit opinions relative to unqualified audit opinions. In the context of this study, qualifications related to GAAP departures (i.e., regarding "material" misstatements) are considered.

PART (A)

Please identify the likelihood which you believe to be associated with each of the following statements. Please circle the appropriate number associated with the likelihood which is most consistent with your views.

If you would prefer to think of the likelihood statements used in this section in terms of probabilities, you may use the following guidelines:

<table>
<thead>
<tr>
<th>Likelihood Statements:</th>
<th>Corresponding Probabilities:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. practically no chance that this would occur;</td>
<td>0-10%</td>
</tr>
<tr>
<td>2. very unlikely that this would occur;</td>
<td>11-25%</td>
</tr>
<tr>
<td>3. somewhat unlikely that this would occur;</td>
<td>26-45%</td>
</tr>
<tr>
<td>4. there is about a 50-50 chance that this would occur;</td>
<td>46-55%</td>
</tr>
<tr>
<td>5. somewhat likely that this would occur;</td>
<td>56-75%</td>
</tr>
<tr>
<td>6. very likely that this would occur;</td>
<td>76-90%</td>
</tr>
<tr>
<td>7. this almost certainly would occur;</td>
<td>91-100%</td>
</tr>
</tbody>
</table>

Consider a client whose financial statements do not actually contain any "material" misstatements. For me to conclude that there are "material" misstatements, and issue a qualified opinion (perhaps inadvertently by error--e.g., due to judgmental areas such as warranties or sales returns, or by making extensions of statistical samples, etc.) would:
e. damage my own standing within my accounting firm

1. practically no chance that this would occur;
2. very unlikely that this would occur;
3. somewhat unlikely that this would occur;
4. there is about a 50-50 chance that this would occur;
5. somewhat likely that this would occur;
6. very likely that this would occur;
7. this almost certainly would occur.

f. demonstrate auditing negligence

1. practically no chance that this would occur;
2. very unlikely that this would occur;
3. somewhat unlikely that this would occur;
4. there is about a 50-50 chance that this would occur;
5. somewhat likely that this would occur;
6. very likely that this would occur;
7. this almost certainly would occur.

g. damage the reputation of the client within the business community

1. practically no chance that this would occur;
2. very unlikely that this would occur;
3. somewhat unlikely that this would occur;
4. there is about a 50-50 chance that this would occur;
5. somewhat likely that this would occur;
6. very likely that this would occur;
7. this almost certainly would occur.

Consider a client whose financial statements contain material misstatements. For me to conclude that there are no "material" misstatements, and issue an unqualified opinion (i.e., not qualified for a GAAP departure; perhaps inadvertently by error--e.g., due to judgmental areas such as warranties or sales returns, or by making extensions of statistical samples, etc.), would:

a. result in litigation by affected parties against my accounting firm

1. practically no chance that this would occur;
2. very unlikely that this would occur;
3. somewhat unlikely that this would occur;
4. there is about a 50-50 chance that this would occur;
5. somewhat likely that this would occur;
6. very likely that this would occur;
7. this almost certainly would occur.
f. result in deteriorated relations with the client
   1. practically no chance that this would occur;
   2. very unlikely that this would occur;
   3. somewhat unlikely that this would occur;
   4. there is about a 50-50 chance that this would occur;
   5. somewhat likely that this would occur;
   6. very likely that this would occur;
   7. this almost certainly would occur.

  g. damage the reputation of the client within the business community
   1. practically no chance that this would occur;
   2. very unlikely that this would occur;
   3. somewhat unlikely that this would occur;
   4. there is about a 50-50 chance that this would occur;
   5. somewhat likely that this would occur;
   6. very likely that this would occur;
   7. this almost certainly would occur.

Consider a client whose financial statements contain "material" misstatements. For me to conclude that there are "material" misstatements and issue an opinion qualified for a GAAP departure would:

a. alert financial statement users to a reporting problem and, thereby, increase the credibility of the audit report to the public
   1. practically no chance that this would occur;
   2. very unlikely that this would occur;
   3. somewhat unlikely that this would occur;
   4. there is about a 50-50 chance that this would occur;
   5. somewhat likely that this would occur;
   6. very likely that this would occur;
   7. this almost certainly would occur.

b. enhance my accounting firm's reputation within the business community
   1. practically no chance that this would occur;
   2. very unlikely that this would occur;
   3. somewhat unlikely that this would occur;
   4. there is about a 50-50 chance that this would occur;
   5. somewhat likely that this would occur;
   6. very likely that this would occur;
   7. this almost certainly would occur.
g. avoid litigation against my accounting firm

1. practically no chance that this would occur;
2. very unlikely that this would occur;
3. somewhat unlikely that this would occur;
4. there is about a 50-50 chance that this would occur;
5. somewhat likely that this would occur;
6. very likely that this would occur;
7. this almost certainly would occur.

Consider a client whose financial statements do not contain any "material" misstatements. For me to conclude that there are no "material" misstatements and issue an unqualified opinion (i.e., not qualified for a GAAP departure) would:

a. fulfill my accounting firm's responsibility to the public

1. practically no chance that this would occur;
2. very unlikely that this would occur;
3. somewhat unlikely that this would occur;
4. there is about a 50-50 chance that this would occur;
5. somewhat likely that this would occur;
6. very likely that this would occur;
7. this almost certainly would occur.

b. result in stronger relations with the client

1. practically no chance that this would occur;
2. very unlikely that this would occur;
3. somewhat unlikely that this would occur;
4. there is about a 50-50 chance that this would occur;
5. somewhat likely that this would occur;
6. very likely that this would occur;
7. this almost certainly would occur.

c. enhance my accounting firm's reputation within the business community

1. practically no chance that this would occur;
2. very unlikely that this would occur;
3. somewhat unlikely that this would occur;
4. there is about a 50-50 chance that this would occur;
5. somewhat likely that this would occur;
6. very likely that this would occur;
7. this almost certainly would occur.
PART (B)

For each of the following statements, please circle one number which most accurately reflects your feelings about the statement.

a. My accounting firm losing a client is
   1. neither bad nor good (indifferent)
   2. slightly bad
   3. moderately bad
   4. bad
   5. extremely bad

b. Litigation against my accounting firm by a client is
   1. neither bad nor good (indifferent)
   2. slightly bad
   3. moderately bad
   4. bad
   5. extremely bad

c. Damage to my accounting firm's reputation within the business community is
   1. neither bad nor good (indifferent)
   2. slightly bad
   3. moderately bad
   4. bad
   5. extremely bad

d. Causing deteriorated relations with a client is
   1. neither bad nor good (indifferent)
   2. slightly bad
   3. moderately bad
   4. bad
   5. extremely bad

e. Damaging my own standing within my accounting firm is
   1. neither bad nor good (indifferent)
   2. slightly bad
   3. moderately bad
   4. bad
   5. extremely bad
k. Enhancing my accounting firm's reputation within the business community is

1. neither bad nor good (indifferent)
2. slightly good
3. moderately good
4. good
5. extremely good

l. Enhancing my own standing within my accounting firm is

1. neither bad nor good (indifferent)
2. slightly good
3. moderately good
4. good
5. extremely good

m. Causing stronger relations with the client is

1. neither bad nor good (indifferent)
2. slightly good
3. moderately good
4. good
5. extremely good

n. Avoiding the undue concern by financial statement users that might occur is a qualified opinion were issued is

1. neither bad nor good (indifferent)
2. slightly good
3. moderately good
4. good
5. extremely good

o. Avoiding litigation against my accounting firm is

1. neither bad nor good (indifferent)
2. slightly good
3. moderately good
4. good
5. extremely good

p. Avoiding the loss of a client is

1. neither bad nor good (indifferent)
2. slightly good
3. moderately good
4. good
5. extremely good
APPENDIX C
SURVEY QUESTIONNAIRE

Please provide the following information about your own background and opinions.

1. To the nearest half year, how long have you been employed as an auditor? _____ years

2(a). Have you worked for your present firm for your entire auditing career? _____ (yes, no)

2(b). If you answered "no" to #2(a), what firm(s) did you previously work for, and how long (to the nearest half year) did you work for the firm(s) in an auditing capacity? (Please list the most recent prior firm first.)

<table>
<thead>
<tr>
<th>Most recent: Firm name</th>
<th>Office (city)</th>
<th>How long</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Next most recent: Firm name</th>
<th>Office (city)</th>
<th>How long</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Next most recent: Firm name</th>
<th>Office (city)</th>
<th>How long</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Across your entire auditing career, what was the length of time (to the nearest half year) you spent as a staff level auditor, senior, manager, and partner? (Since the hierarchy is somewhat different across firms, please include the time spent as an "assistant" along with "staff" and include time spent as a "supervisor" along with "manager.")

   Staff _____ years
   Senior _____ years
   Manager _____ years
   Partner _____ years
   Total _____ years (should agree to #1 above)
   =====

4. Are you a CPA? _____ (yes, no)

5. Have you received a college/university degree? _____ (yes, no)
   [If you answered "no" to #5, please skip to #8(a).]
12(a). Do you believe that, under the same circumstances, different auditors should establish the same materiality threshold at the **planning** stage? (Circle one number)

1. Strongly yes;
2. Moderately yes;
3. Neutral (neither yes nor no);
4. Moderately no;
5. Strongly no;

12(b). Do you believe that, under the same circumstances, different auditors should establish the same materiality threshold at the **evaluation** stage? (Circle one number)

1. Strongly yes;
2. Moderately yes;
3. Neutral (neither yes nor no);
4. Moderately no;
5. Strongly no;

13(a). Do you believe that more detailed guidelines are needed to assist auditors in quantifying their materiality judgments (in order to achieve a greater degree of consensus across the profession) at the **planning** stage of the audit (termed the "preliminary judgment about materiality" per SAS No. 47)? (Circle one number)

1. Strongly yes;
2. Moderately yes;
3. Neutral (neither yes nor no);
4. Moderately no;
5. Strongly no;

13(b). Do you believe that more detailed guidelines are needed to assist auditors in quantifying their materiality judgments (in order to achieve a greater degree of consensus across the profession) at the **evaluation** stage of the audit (i.e., the "final" materiality judgment)? (Circle one number)

1. Strongly yes;
2. Moderately yes;
3. Neutral (neither yes nor no);
4. Moderately no;
5. Strongly no;
17(a). Compared with other auditors within your firm, in similar engagements, do you believe you tend to perform more or less extensive audit tests? (Circle one number)

1. Much more extensive than most;
2. Somewhat more extensive than most;
3. About the same as most;
4. Somewhat less extensive than most;
5. Much less extensive than most;
6. No opinion--cannot say;

17(b). Compared with other auditors from other firms, in similar engagements, do you believe you tend to perform more or less extensive audit tests? (Circle one number)

1. Much more extensive than most;
2. Somewhat more extensive than most;
3. About the same as most;
4. Somewhat less extensive than most;
5. Much less extensive than most;
6. No opinion--cannot say;

18(a). Compared with other auditors within your firm, do you believe you tend to spend relatively more or less time related to audit planning activities? (Circle one number)

1. Much more time than most;
2. Somewhat more time than most;
3. About the same as most;
4. Somewhat less time than most;
5. Much less time than most;
6. No opinion--cannot say;

18(b). Compared with other auditors from other firms, do you believe you tend to spend relatively more or less time related to audit planning activities? (Circle one number)

1. Much more time than most;
2. Somewhat more time than most;
3. About the same as most;
4. Somewhat less time than most;
5. Much less time than most;
6. No opinion--cannot say;
24(a). Would you have found additional information helpful in estimating the audit hours for the selected cases? _____ (yes, no)

24(b). If you answered "yes" to #24(a), please specify.

____________________________________________________________________________________
____________________________________________________________________________________
____________________________________________________________________________________

25(a). Please allocate 100 points to the three factors manipulated in the 28 cases according to how important each was in arriving at your materiality judgments regarding the income statement (i.e., important factors should be assigned more points than the less important factors).

a. Total assets ______
b. Net sales ______
c. Income (loss) before taxes ______
d. Other (e.g., some combination of factors) ______

Total 100

25(b). Please allocate 100 points to the three factors manipulated in the 28 cases according to how important each was in arriving at your materiality judgments regarding the balance sheet (i.e., important factors should be assigned more points than the less important factors).

a. Total assets ______
b. Net sales ______
c. Income (loss) before taxes ______
d. Other (e.g., some combination of factors) ______

Total 100

25(c). Please allocate 100 points to the three factors manipulated in the 28 cases according to how important each was in arriving at the estimated audit hours (i.e., important factors should be assigned more points than the less important factors).

a. Total assets ______
b. Net sales ______
c. Income (loss) before taxes ______
d. Other (e.g., some combination of factors) ______

Total 100
APPENDIX D
RESULTS OF ANALYSES (BALANCE SHEET JUDGMENTS)

Introduction

As pointed out previously, SAS No. 47 suggests that auditors should consider their preliminary judgments about materiality at the aggregate financial statement level. The SAS indicates (paragraph 12) that these judgments may be made appropriately from either an income statement or balance sheet perspective. In general, prior research has scarcely examined auditors' planning stage materiality judgments. In particular, there has been no distinction made in the academic or professional literatures between materiality judgments from an income statement versus balance sheet point of view.

In this study, the income statement materiality judgments are deemed to be the dependent variable of primary interest. Virtually all of the prior studies of auditors' materiality judgments (nearly all of which focused on the evaluation stage) have shown income considerations to dominate the judgments. Consequently, it was expected that income (loss) before taxes would be the primary cue influencing auditors' judgments in this experiment. Along the same line, income statement considerations are generally believed to drive auditors' planning stage materiality judgments, in practice (e.g., Zuber, et al. [1983, p. 43]).

283
These discussions are followed by an examination of the participants' individual ANOVA models for the balance sheet judgments. After that, test-retest consistency and order effects are considered.

**Level of Responsibility Effects**

As a starting point, subjects' responses were standardized on a case-by-case basis, across all 72 subjects. For each individual case, the responses of all the subjects on that case were transformed to a distribution with mean zero and variance one. (Thus, an individual with an above average response on that case would have a positive standardized score for that case; an individual with a below average response on that case would have a negative standardized score for that case.) These standardized scores were summed for each participant across all 24 cases to obtain an overall measure of the extremeness of each subject's balance sheet judgments. Figure 10 plots these standardized score, by level of responsibility, for the seniors, managers, and partners (levels 2-4, respectively).

**Location**

Figure 10 suggests that there are no systematic differences among the seniors, managers, and partners in their balance sheet materiality judgments with respect to location. Further, no outliers can be identified. As described in Chapters III and IV, this study
investigates two distinct potential sources of systematic differences among the participants (i.e., experience-related and firm-related effects). To simultaneously consider both possible effects in the analysis of effects associated with level of responsibility, the analyses for differences among the seniors, managers, and partners were performed on a firm-by-firm basis.

Therefore, six separate one-way ANOVAs were conducted to identify whether there were systematic differences in location across the three levels of responsibility. As summarized in Table 20, level of responsibility was not significant for any of the firms at any reasonable level. Fisher's method for combining tests was used to derive an overall test statistic regarding differences associated with level of responsibility (Hedges and Olkin [1985, p. 37]). The resulting test statistic was approximately 8.87 and was not significant at any reasonable level (the critical value for the .05 level was 21.03). Thus, there were no real differences in location among the three levels of responsibility.

One further type of comparison was made to gain possible insights regarding differences associated with level of responsibility. Effects were examined within each of the two audit technologies. (Firms #1-3 were the relatively structured firms and firms #4-6 were the relatively unstructured firms.) Two separate one-way ANOVAs were run to identify whether there were differences among the three levels of responsibility for either audit technology group. Again,
level of responsibility was not significant at the .10 level for either of the two audit technologies.

Variation

Figure 10 does not portray any conspicuous differences in variability among the seniors, managers, and partners. The following sums of squares were associated with each level of responsibility:

<table>
<thead>
<tr>
<th></th>
<th>Seniors</th>
<th>9,832</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managers</td>
<td>9,521</td>
<td></td>
</tr>
<tr>
<td>Partners</td>
<td>10,630</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>29,983</td>
<td></td>
</tr>
</tbody>
</table>

Effects associated with level of responsibility with respect to variation were examined within each of the six firms. Six separate applications of Cochran's test for homogeneity of variance were conducted, using the sums of squares for each level of responsibility for each firm, to identify whether there were any significant differences among the levels of responsibility for any of the firms. Using the method described by Cochran [1941], the specific p-values were inferred with respect to level of responsibility for each firm. These p-values were then combined into an overall test statistic using Fisher's method for combining tests.

No differences associated with level of responsibility were found for any of the firms at the .05 level. When the six p-values were combined using Fisher's method, an overall test statistic of approximately 11.09 resulted. This was not significant at any
Table 21

Summary of Level of Responsibility Effects
For the Balance Sheet Judgments

<table>
<thead>
<tr>
<th>Comparisons Made</th>
<th>Re: Location</th>
<th>Re: Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Level of responsibility, by firm affiliation</td>
<td>No significant differences were found at the .10 level.</td>
<td>No significant differences at the .05 level.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Overall effect associated with level of responsibility</td>
<td>No differences among the seniors, managers, and partners at the .10 level.</td>
<td>No significant differences at the .05 level.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Level of responsibility, by audit technology</td>
<td>No significant differences for either audit technology at the .10 level.</td>
<td>No significant differences at the .05 level.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: This table summarizes the analyses identified in Table 4. Table 4 summarizes the primary method of analysis for differences as to location and variation, identifies the data on which the above conclusions are based, and ties the analyses conducted to the notation introduced in Figure 7.
Figure 11

Standardized Scores for the Balance Sheet Materiality Judgments
Plotted by Firm Affiliation
firm #5 had a higher balance sheet materiality threshold than four of the other firms (p ≤ .003 where .003 = .05/15, since there were 15 between-firm comparisons). There was a real difference in central tendency among the firms with respect to their balance sheet materiality judgments.

Two types of supplemental comparisons were made to gain insights about the overall firm differences identified above. First, firm differences were examined, separately, across each of the three levels of responsibility. Second, firm differences were examined separately within each of the two audit technologies.

Three separate one-way ANOVAs were conducted to identify whether there were any significant differences as to location across the firms at each level of responsibility. Firm effects were highly significant for the seniors (p = .006). Tukey's multiple comparison test identified the seniors from firm #5 to be different from the seniors from firms #4 and 6. Firm affiliation was not significant for the managers (p = .06) or partners (p = .37), although the managers could be viewed as borderline. Despite the non-significance among the managers and partners, an apparently modest firm effect was associated with them, too. The managers from firm #5 had the highest standardized scores, followed by the managers from firm #6; the managers from firm #4 had the lowest scores. The partners from firm #6 had the highest scores, followed by the partners from firm #5; the partners from firm #4 had the lowest scores. As a practical matter, at each level of
squares associated with firm #6, the most variable firm, Cochran's test was applied to the remaining sums of squares. No difference could be demonstrated among the remaining firms at the .05 level.

To clarify which particular firms differed (and because Cochran's test for homogeneity of variance may be affected by non-normal data), an additional test of variation was conducted. Ratios of the sums of squares (i.e., F-tests) were calculated for each pair of firms for all pairwise comparisons (there were fifteen such comparisons). Significant differences were identified using the conservative Bonferonni simultaneous confidence intervals at the .05 level. Only one such comparison was significant; firm #6 was significantly higher in variation than firm #1. That result is consistent with Cochran's test and clarifies which firms differ.

Similar to the tests regarding location reported above, two types of supplemental comparisons were made to gain insights about the overall firm differences as to variation. First, the variances were examined across the firms, separately, by level of responsibility. These variances are summarized in Table 23. Second, firm differences were examined, separately within each of the audit technology groups.

Cochran's test for homogeneity of variance was applied across firms, separately for each level of responsibility, based on the sums of squares associated with each level within each firm. The test was applied sequentially, deleting the most variable firm when significance was observed and applying the test to the remaining
firms. For the seniors, firm #6 was the most variable, followed by firms #3 and 2. For the managers, firm #5 was the most variable, followed by firm #3. Firm #6 was the most variable for the partners (p ≤ .05). These were the only significant firm differences, by level, at the .05 level. Firm differences were fairly pervasive across levels of responsibility.

To check for firm differences as to variation within each audit technology group, two separate applications of Cochran's test for homogeneity of variance were made. The sums of squares for the firms were used (the same as those reported above with respect to overall firm effects), but grouped by audit technology. No differences were found among the firms within either audit technology group at the .05 level.

Summary of Firm Effects

Differences were identified at the firm level with respect to both location and variation. These results are summarized in Table 24.

Systematic differences were identified among the firms with respect to location. In particular, the participants from firm #5 had higher balance sheet materiality thresholds than the participants from each of the other firms. The extremeness of the judgments involving firm #5 was further shown by an analysis of firm affiliation, level-by-level. On average, the seniors, managers, and partners from firm
#5 all appeared to have high balance sheet judgments relative to their counterparts from other firms (although only the seniors could be identified as significant at the .05 level). The greatest firm differences with respect to location occurred between two unstructured firms--firm #5 had the highest standardized balance sheet materiality scores and firm #4 had the lowest. No firm differences as to location were detected among the structured firms at the .05 level.

Firm differences were also identified with respect to variability. Firm #6 was shown to be significantly more variable than firm #1; no other differences were significant at the .05 level. Firm differences were identified at each level of responsibility. Among the seniors and among the partners, firm #6 was the most variable; among the managers, firm #5 was the most variable. Within each audit technology group, however, no firm differences could be identified. None of the structured firms was distinctly more variable than the others; similarly, none of the unstructured firms was distinctly more variable than the others.

Audit Technology Effects

Based on examination of the standardized materiality scores, plotted by firm affiliation in Figure 11 and the analyses of firm effects reported above, audit technology is apparently not related to either the central tendency or variability of the firms' balance sheet
<table>
<thead>
<tr>
<th>Source</th>
<th>Degrees of Freedom</th>
<th>Sums of Squares</th>
<th>Mean Square</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1</td>
<td>870.4</td>
<td>870.4</td>
<td>2.05</td>
<td>.16</td>
</tr>
<tr>
<td>Residual</td>
<td>70</td>
<td>29,685.8</td>
<td>424.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>71</td>
<td>30,556.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Moreover, as pointed out above, only one of the fifteen F-tests (based on all pairwise comparisons among the firms' sums of squares) was significant. Nine of the fifteen comparisons involved firms having different audit technologies, and only the comparison between firms #1 and 6 was significant. Consequently, the evidence of differences in variation between the two audit technologies is ambiguous.

A supplemental analysis examined each level of responsibility separately for differences as to variation between the structured and unstructured firms. The sums of squares were calculated for each level of responsibility within each of the two audit technology groups. Three separate applications of Cochran's test for homogeneity of variance were made, one for each level. Although the sums of squares associated with each level were larger for the subjects from unstructured firms, none of the levels of responsibility was significantly different between the two audit technology groups at the .05 level.

Summary of Audit Technology Effects

No difference was identified as to location between the structured and unstructured firms. However, with respect to variation, a significant difference was observed; as pointed out, this result should be viewed somewhat cautiously. These results are summarized in Table 26.
With respect to location, no overall differences were identified between the structured and unstructured firms. Similarly, no differences were found between the two audit technology groups, on a level-by-level basis, for the seniors, managers, and partners. The notion of audit technology was not helpful in understanding the differences observed at the firm level.

With respect to variation, an overall difference was found between the structured and unstructured firms; the unstructured firms, as a group, had greater variability based on the sums of squares associated with the three firms in each technology group. However, only two of the three unstructured firms were visibly more variable than the structured firms. Furthermore, only one of the nine F-tests (involving all possible pairwise comparisons of the firms' sums of squares) for firms having different audit technologies was significant. Therefore, the evidence of audit technology differences as to variation must be viewed as ambiguous. No differences were found between the structured and unstructured firms, on a level-by-level basis, for the seniors, managers, and partners.

Additional Analyses

Similar to the discussion of the income statement materiality judgments in Chapter IV, this section reports the results of several additional types of analyses with respect to the balance sheet
separate financial statements on many or all of the cases). Two-way interactions were seldom significant; the most frequently significant interaction was income x assets, which was significant for only six subjects. Table 27 summarizes the significant cues (p ≤ .05), by level of responsibility and, separately, by firm affiliation.

Table 28 summarizes the significant cues for the subjects for whom only one variable was significant. Only one cue was significant for 61 of the 72 subjects. Of these, total assets was the sole significant cue for 56 of the subjects. The individual ANOVA models of subjects' balance sheet materiality judgments do not suggest systematic differences among the subjects, either with respect to level of responsibility or firm affiliation. Total assets was the dominant cue for the preponderance of subjects. Compared to the discussion of the income statement judgments in Chapter IV, individual differences were less conspicuous with respect to the balance sheet judgments.

Test-retest Consistency

In order to provide an indication of the meaningfulness of subjects' responses, four cases were repeated for each participant after the 24 original cases were presented. The same cases were repeated for each subject. As described in Chapter IV, the Pearson product-moment correlation coefficient was calculated across all 72 subjects. Secondly, paired t-tests were calculated, by case, to
Table 28

Only One Significant Cue (p < .05)
For the Balance Sheet Judgments
By Level of Responsibility, By Firm Affiliation

By Level of Responsibility:

<table>
<thead>
<tr>
<th></th>
<th>Seniors</th>
<th>Managers</th>
<th>Partners</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income (loss) before taxes</td>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Net sales</td>
<td>2</td>
<td></td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Total assets</td>
<td>16</td>
<td>22</td>
<td>18</td>
<td>56</td>
</tr>
</tbody>
</table>

By Firm Affiliation:

<table>
<thead>
<tr>
<th></th>
<th>Structured</th>
<th>Unstructured</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Firm #1</td>
<td>#2</td>
<td>#3</td>
</tr>
<tr>
<td>Income (loss) before taxes</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Net sales</td>
<td>-</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Total assets</td>
<td>11</td>
<td>8</td>
<td>7</td>
</tr>
</tbody>
</table>
able and willing to provide meaningful (i.e., consistent) responses to the experiment's materials.

Order Effects

The 24 cases were presented to the subjects in two different randomized orders. The responses to the balance sheet judgments were examined to identify whether the order of presentation significantly affected subjects' responses in terms of location and/or variation. If order of presentation influenced subjects' judgments, their responses could be regarded as artifacts of the experiment.

No differences were detected between the two orders of presentation, either as to location or variation. A one-way ANOVA of the standardized materiality scores indicated no significant difference in means between the two orders (p=.82). Cochran's test for homogeneity of variance did not indicate any difference as to variation between the two orders at the .05 level, using the sums of squares calculated with respect to each order. Apparently, the participants' judgments were not influenced by the order in which the cases were presented.


Chandler, R. "Materiality--Does It Need To Be a Guessing Game?" Accountancy, (February, 1985), pp. 84-86.


