



LIBRARY
Michigan State
University



This is to certify that the
dissertation entitled
INFORMATION INTEGRATION
IN AUDIT PLANNING

presented by
David Samuel Kerr

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Accounting

D. Dewey Ward
Major professor

Date August 4, 1989

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.

DATE DUE	DATE DUE	DATE DUE
REAN 05 1993		
SEP 27 1993		
262		

MSU Is An Affirmative Action/Equal Opportunity Institution

INFORMATION INTEGRATION
IN AUDIT PLANNING

By

David Samuel Kerr

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Accounting

1989

6040202

ABSTRACT

INFORMATION INTEGRATION IN AUDIT PLANNING

By

David Samuel Kerr

Professional audit judgment is the key component in audit planning. AICPA publications and firms' audit manuals provide *general* guidelines on the planning of audits, but the actual planning of individual audits is left to the auditor's professional judgment. Very little is known about the actual judgment processes involved in audit planning. For instance, while the accounting literature suggests that various types of information should be considered when planning an audit, the judgment processes involved in the use of this information when making decisions is not well understood.

This research examines the following issues: (1) the manner in which auditors utilize and integrate information concerning: (i) the client's internal operating environment, (ii) the effectiveness of the internal control structure, (iii) the outcome of analytical procedures, and (iv) the volatility of the client's external business environment, when forming judgments regarding the appropriate extent of tests of details of balances; (2) the impact of professional audit experience on auditors' audit planning judgments, including a study of the relationship between professional experience and judgmental agreement between auditors, an analysis of sources of judgmental agreement between auditors, and how and where professional expertise relating to audit program planning is developed; and (3) the manner in which auditors integrate results of tests of individual internal controls into a single

judgment of the effectiveness of the internal control structure.

An experiment was performed to investigate the issues mentioned above. Results of the experiment indicate a high degree of configurality in auditors' extent-of-testing judgments. This was revealed by several significant two-way interactions and a significant three-way interaction involving the client's internal operating environment, internal control structure, and results of analytical procedures.

Results also indicate that professional audit experience plays an important role in audit planning. An analysis of the sources of judgmental agreement between auditors at different levels of experience revealed that audit experience also plays an important role in increasing cue-weighting agreement between auditors, but does not significantly affect judgmental consistency or agreement on the appropriate degree of configural cue utilization.

Copyright by
DAVID SAMUEL KERR
1989

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to the members of my dissertation committee -- Dewey Ward, chairperson; Alvin Arens; and Frank Boster -- for their guidance, encouragement, and wisdom. Working with each of them has been a truly enjoyable and enlightening experience.

I would also like to express my appreciation to the anonymous auditors who generously offered their time and talents to participate in this research. Without their assistance, this research would not have been possible.

I wish to also thank my parents, Glen and Rachel Kerr, for the guidance they have offered to me throughout my life, and for their encouraging phone calls and support during my time at Michigan State University.

Finally, my utmost gratitude and appreciation goes to my wonderful wife, Elva, who, through her continual encouragement and patience, has been a source of great strength and joy.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	x
CHAPTER 1 - INTRODUCTION	1
1.1 AUDIT PLANNING AND AUDIT RISK.	2
1.1.1 Reliance on the Internal Operating Environment.	5
1.1.2 Reliance on the Internal Control Structure.	6
1.1.3 Reliance on Analytical Procedures	9
1.1.4 Business Risk, Acceptable Audit Risk, and the External Environment.	10
1.2 MOTIVATION OF THE RESEARCH	12
1.2.1 Restrictions in Scope of Previous Research on Audit Program Planning.	13
1.3 SCOPE OF THE RESEARCH.	15
1.3.1 Stages in the Audit Planning Judgment	15
1.4 OBJECTIVES OF THE RESEARCH	18
1.5 ORGANIZATION OF THE DISSERTATION	20
CHAPTER 2 - REVIEW OF RELATED RESEARCH AND DEVELOPMENT OF HYPOTHESES	21
2.1 AUDIT PLANNING AND THE INTERNAL OPERATING ENVIRONMENT	22
2.2 AUDIT PLANNING AND THE INTERNAL CONTROL STRUCTURE.	28
2.3 AUDIT PLANNING AND ANALYTICAL PROCEDURES	30
2.4 AUDIT PLANNING AND THE CLIENT'S EXTERNAL BUSINESS ENVIRONMENT	38
2.5 INTERNAL CONTROLS AND THE INHERENT LIKELIHOOD OF ERRORS	41
2.6 INTERNAL CONTROLS AND ANALYTICAL PROCEDURES.	43
2.7 ANALYTICAL PROCEDURES AND THE EXTERNAL BUSINESS ENVIRONMENT	44
2.8 EFFECTS OF PROFESSIONAL EXPERIENCE AND TASK COMPLEXITY ON AUDIT PLANNING	45
CHAPTER 3 - METHODOLOGY	59
3.1 INTRODUCTION	59
3.2 RESEARCH INSTRUMENT AND ADMINISTRATION OF THE EXPERIMENT	60
3.3 COMPONENTS-INTEGRATION STAGE AUDIT JUDGMENTS	63
3.3.1 Tests of Summing Versus Averaging Integration of Audit Evidence	65
3.3.2 Evaluation of Internal Controls -- Independent Variables	67

3.3.3	Evaluation of Internal Controls -- Dependent Variable	68
3.4	FACTORS-INTEGRATION STAGE AUDIT JUDGMENTS	69
3.4.1	Independent Variables	69
3.4.2	Dependent Variables	71
3.4.3	Tests of Hypotheses	73
3.5	EFFECTS OF EXPERIENCE AND AUDIT TASK	78
3.6	PILOT TESTS.	82
3.7	SUBJECTS	83
CHAPTER 4	- RESULTS	85
4.1	INTRODUCTION	85
4.2	USE AND INTEGRATION OF EVIDENCE IN AUDIT PLANNING	86
4.2.1	Manipulation Checks	86
4.2.2	Budgeted Audit Hours	87
4.2.3	Extent-of-Testing Ratings.	102
4.2.4	Weighting of Evidence	124
4.3	EFFECTS OF EXPERIENCE.	127
4.3.1	Judgmental Agreement Based on Correlations	127
4.3.2	Judgmental Agreement Based on Mean Absolute Differences.	129
4.3.3	Decomposition of Consensus	134
4.3.4	Decomposition of Experience Levels	138
4.3.5	G, C, & R for Seniors and Managers	145
4.4	AVERAGING- VERSUS SUMMING-TYPE EVIDENCE INTEGRATION	151
4.5	ADDITIONAL ANALYSES.	156
CHAPTER 5	- SUMMARY, CONCLUSIONS, AND SUGGESTIONS FOR FUTURE RESEARCH	162
5.1	SUMMARY OF EXPERIMENTAL FINDINGS	162
5.1.1	Budgeted Audit Hours Judgments	162
5.1.2	Extent-of-Testing Ratings.	166
5.1.3	Weighting of Evidence.	171
5.1.4	Effects of Professional Experience	173
5.1.5	Decomposition of Consensus	175
5.1.6	Decomposition of Experience Levels	175
5.1.7	Averaging- Versus Summing-Type Evidence Integration	177
5.2	LIMITATIONS OF THE STUDY	178
5.3	SUGGESTIONS FOR FUTURE RESEARCH	179
LIST OF REFERENCES		182
APPENDICES		
A.	Booklet 1 - Volatile External Environment Condition	188
B.	Booklet 2 (Partial) - Sample Audit Planning Case.	204
C.	Booklet 3 (Partial) - Sample Internal Control Evaluation Case and Debriefing Questionnaire	208

LIST OF TABLES

<u>Table</u>	<u>Title</u>	<u>Page</u>
3.1	Number of Subjects at each Experience Level	84
4.1	Budgeted Audit Hours - Cell Means and Standard Deviations	89
4.2	Budgeted Audit Hours - ANOVA Table	91
4.3	Budgeted Audit Hours - Experience Level x Internal Operating Environment (A x C) Interaction Contrasts . . .	97
4.4	Budgeted Audit Hours - Experience Level x Internal Control (A x D) Interaction Contrasts	98
4.5	Budgeted Audit Hours - Experience Level x Results of Analytical Procedures (A x E) Interaction Contrasts. .	99
4.6	Extent-of-Testing Ratings - Cell Means and Standard Deviations	103
4.7	Extent-of-Testing Ratings - ANOVA Table	105
4.8	Extent-of-Testing Ratings - Analysis of Simple Interactions	109
4.9	Extent-of-Testing Ratings - Eta-Squared Indices - Students, Seniors, and Managers	125
4.10	Inter-Auditor Consensus as Measured by Pair-Wise Correlation Coefficients	127
4.11	Inter-Auditor Consensus Based on Correlation Coefficients - Students & Practitioners - ANOVA Table .	129
4.12	Inter-Auditor Consensus as Measured by Mean Absolute Differences (CNS_{MAD})	130
4.13	Inter-Auditor Consensus Based on Mean Absolute Differences (CNS_{MAD}) - Students & Practitioners - ANOVA Table	132
4.14	Mean Values for G_{MAD} , R_{MAD} , and C_{MAD}	135

4.15	Cue Weighting Agreement (G_{MAD}) - Students & Practitioners - ANOVA Table	135
4.16	Judgment Consistency (R_{MAD}) - Students & Practitioners - ANOVA Table	136
4.17	Configural Cue Utilization Agreement (C_{MAD}) - Students & Practitioners - ANOVA Table	137
4.18	Inter-Auditor Consensus as Measured by Mean Absolute Differences (CNS_{MAD}) - Managers, Seniors, & Students	138
4.19	Inter-Auditor Consensus (CNS_{MAD}) - Students and Seniors - ANOVA Table	142
4.20	Inter-Auditor Consensus (CNS_{MAD}) - Seniors and Managers - ANOVA Table	143
4.21	Mean Values for G_{MAD} , R_{MAD} , and C_{MAD} - Managers, Seniors, & Students	145
4.22	Cue-Weighting Agreement (G_{MAD}) - Students and Seniors - ANOVA Table	146
4.23	Judgment Consistency (R_{MAD}) - Students and Seniors - ANOVA Table	147
4.24	Configural Cue Utilization Agreement (C_{MAD}) - Students and Seniors - ANOVA Table	148
4.25	Cue-Weighting Agreement (G_{MAD}) - Seniors and Managers - ANOVA Table	149
4.26	Judgment Consistency (R_{MAD}) - Seniors and Managers - ANOVA Table	149
4.27	Configural Cue Utilization Agreement (C_{MAD}) - Seniors and Managers - ANOVA Table	149
4.28	Internal Control Evaluation Ratings - Cell Means and Standard Deviations	153
4.29	Internal Control Evaluation Ratings - ANOVA Table	154

LIST OF FIGURES

<u>Figure</u>	<u>Title</u>	<u>Page</u>
1.1	Stages of Extent-of-Tests-of-Details Judgment	17
2.1	Summary of Research Hypotheses	56
3.1	Graphical Test of Summing vs. Averaging Information Integration Processes	66
3.2	Summary of Variables	73
4.1	Budgeted Audit Hours - Interaction Between Experience Level and Internal Operating Environment (A x C).	93
4.2	Budgeted Audit Hours - Interaction Between Experience Level and Internal Control (A x D)	94
4.3	Budgeted Audit Hours - Interaction Between Experience Level and Results of Analytical Procedures (A x E).	95
4.4	Extent-of-Testing Ratings - Interaction Between Experience Level and Internal Operating Environment (A x C)	106
4.5	Extent-of-Testing Ratings - Three-Way Interaction Between Internal Operating Environment, Internal Control, and Analytical Procedures (C x D x E)	108
4.6	Extent-of-Testing Ratings - Analysis of C x D x E Interaction	111
4.7	Hypothesized Interaction Between Internal Operating Environment (C) and Internal Control (D)	117
4.8	Extent-of-Testing Ratings - Three-Way Interaction Between Internal Operating Environment, Internal Control, and Analytical Procedures (C x D x E).	121
4.9	Experience Level x Audit Task Interaction (CNS _{MAD}) - Students and Practitioners	132
4.10	Experience Level x Audit Task Interaction (G _{MAD}) - Students and Practitioners	136

4.11	Experience Level x Audit Task Interaction (CNS_{MAD}) - Students and Seniors	142
4.12	Experience Level x Audit Task Interaction (CNS_{MAD}) - Seniors and Managers	144
4.13	Experience Level x Audit Task Interaction (G_{MAD}) - Students and Seniors.	147
4.14	Steps in Identifying Summing vs. Averaging Integration Processes	152
4.15	Internal Control Evaluation Ratings - C x E Interaction	155
4.16	Internal Control Evaluation Ratings - D x E Interaction	155
5.1	Summary of Tests of Hypotheses One Through Seven Based on Budgeted Audit Hours Responses	166
5.2	Summary of Tests of Hypotheses One Through Seven Based on Extent-of-Testing Ratings	172

CHAPTER 1 - INTRODUCTION

The evaluation and integration of items of information is a task which faces experts in virtually every field. The degree of success attained by the expert is often a function of the expert's ability to appropriately evaluate and integrate the information available. In the auditing profession, auditors are frequently required to integrate various items of information and evidence into a single judgment. For instance, the results of tests of individual internal controls are integrated when forming a judgment concerning the effectiveness of the internal control structure. When evaluating the fairness of a reported account balance, the auditor considers several items of information, including the preliminary review of the client's internal control structure, results of tests of the effectiveness of the structure, results of substantive tests of the account balance, level of materiality, and desired audit risk. The relevant factors must be evaluated, weighed, and integrated into a judgment regarding the acceptability of the reported account balance. Further integration of information occurs as the auditor forms judgments regarding the acceptability of the reported values of groups of accounts such as total assets and net income, and the acceptability of the financial statements as a whole.

This research focuses on (1) auditors' judgment processes involved in audit planning, (2) the effects of professional experience on auditors' audit planning and internal control evaluation judgments, and (3) the nature of the integration functions involved in auditors' evaluations of internal control structures. The use and integration of

four types of information which the auditing literature suggests may be relevant to the planning of an audit is examined. These four types of information are: (1) information concerning conditions in the client's internal operating environment, (2) results of tests of the internal control structure, (3) results of analytical procedures, and (4) information concerning the client's external business environment. The effects of professional experience on auditors' use of the above four types of information is also examined. Finally, the manner in which results of tests of individual internal controls are integrated when evaluating the effectiveness of the internal control structure is examined. In the next section, a brief discussion is offered of the nature of each of the above four types of information.

1.1 AUDIT PLANNING AND AUDIT RISK

When planning an audit, the auditor should take into account the risk that material errors exist in the financial statements of the client and design the audit so as to provide reasonable assurance that all material errors will be detected.

The auditor should consider audit risk and materiality both in (a) planning the audit and designing auditing procedures and (b) evaluating whether the financial statements taken as a whole are presented fairly in conformity with generally accepted accounting principles. The auditor should consider audit risk and materiality in the first circumstance to obtain sufficient competent evidential matter on which to properly evaluate the financial statements in the second circumstance. (AICPA, [1986], AU 312.08)

Audit risk is the risk that the auditor may unknowingly fail to appropriately modify his/her opinion on financial statements that are materially misstated. This risk consists of two parts: (1) the risk, which is uncontrollable by the auditor, that material misstatements will

occur in the client's accounting records and go undetected by the client, and (2) the controllable risk that misstatements will go undetected by the auditor. The auditor may rely on elements of the internal operating environment and on the internal control structure to reduce the first risk, and on analytical procedures and tests of details to reduce the second risk (AICPA [1986], AU 350.08).

The audit risk model provides a finer partitioning of audit risk by identifying four separate components: (1) the risk that material misstatements will occur, assuming no related internal controls (inherent risk); (2) the risk the material misstatements will not be prevented or detected by internal controls (control risk); (3) the risk that the auditor's analytical procedures will not detect the misstatements (analytical procedures risk); and (4) the risk that the auditor's substantive tests of details will not detect the misstatements (test of details risk). The first two components are uncontrollable by the auditor and relate to the risk that material errors exist in the financial statements before audit work is begun. The last two components are controllable by the auditor and represent the risk that the auditor will fail to detect the misstatements. The audit risk model depicts these four components of audit risk as a multiplicative function:

$$AR = IR \times CR \times AnR \times TD$$

A more useful form of the audit risk model for audit planning is one in which an appropriate risk of incorrect acceptance for a substantive test of details (acceptable test of details risk (ATD)) is determined by first specifying the auditor's acceptable audit risk (AR), along with the auditor's subjectively quantified levels of

inherent risk, control risk, and analytical procedures risk. This is known as the planning form of the audit risk model:

$$ATD = \frac{AAR}{IR \times CR \times AnR}$$

An increase in the acceptable level of audit risk would be accompanied by an increase in acceptable test of details risk, while an increase in inherent risk, control risk, or analytical risk would be accompanied by a decrease in acceptable test of details risk. In turn, if the acceptable test of details risk increases (decreases), the auditor would select a larger (smaller) sample for the test of details (AICPA, [1986], AU 350.19).

The auditor reaches his/her acceptable level of audit risk by gathering substantive audit evidence. The quantity of evidence necessary to achieve the acceptable level of risk depends, in part, on the auditors' evaluation of factors which affect the inherent likelihood of misstatements occurring, and on the auditor's evaluations of, and tests of, the internal control structure. Factors which affect the inherent likelihood of misstatements, such as conditions in the client's internal operating environment and characteristics of the account under audit, will influence the auditor's assessment of inherent risk. Likewise, the auditor's evaluation of, and tests of, the internal control structure will influence his/her assessment of control risk. As indicated in the planning form of the audit risk model, changes in the auditor's assessment of inherent risk and control risk influence the acceptable test of details risk, which, in turn, influences the level of evidence needed to achieve the acceptable level of audit risk.

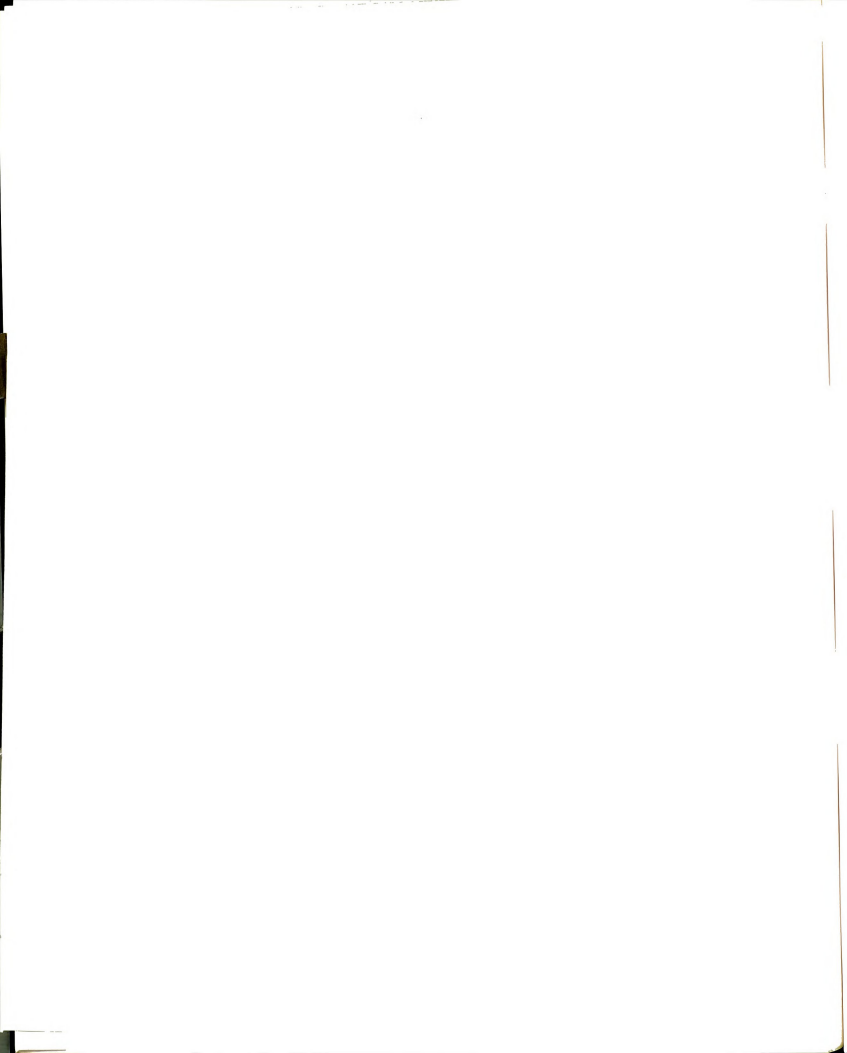
1.1.1 Reliance on the Internal Operating Environment

Information about the likelihood of material misstatements occurring in the financial statements is one type of information that the auditor considers during the course of an audit. When the auditor assesses inherent risk at less than 100%, he/she is relying, to some extent, on characteristics of the internal operating environment and of the account being audited to reduce the likelihood that material misstatements have occurred in that account during the period under audit. In the auditor's evaluation of inherent risk, this likelihood is considered independently of internal controls, since control risk is dealt with separately in the audit risk model. The importance of this evaluation is described in the following statement by Arens and Loebbecke:

The inclusion of inherent risk in the audit risk model is one of the most important concepts in auditing. It implies that auditors should attempt to predict where errors are most and least likely in the financial statement segments. This information affects the total amount of evidence the auditor is required to accumulate and influences how the auditor's efforts to gather the evidence are allocated among the segments of the audit. (Arens & Loebbecke, [1988], pg. 246).

This risk of material misstatements is greater for some accounts than for others. For example, the cash account may be more susceptible to misstatement than the land account since there is likely to be a greater number of transactions affecting the cash account than the land account during any given period. Therefore, the potential for misstatements is greater in the cash account.

In addition to differences in the risk of misstatements for different accounts, the risk of misstatement for a given account may be greater for some clients than for others. Characteristics of a client's



Internal operating environment may play an important role in the risk that misstatements will occur. For example, when the client's management lacks integrity, there is a substantial increase in the likelihood that the financial statements are misstated due to management fraud. When there has been an increase in the turnover rate among employees in the accounting department, the risk of errors occurring is also likely to have increased since new employees may lack the knowledge to correctly handle the recording of transactions. If errors have been found in previous audits of a given account for a particular client, the auditor is likely to assess inherent risk as high since many types of errors tend to recur and the client may not have corrected their causes. On the other hand, when conditions in the client's internal operating environment are favorable (e.g., when the integrity of management is high, when accounting personnel are competent, when previous years' audits revealed no errors in a particular account), the auditor may rely on the internal operating environment to reduce inherent risk. This results in an increase in the acceptable level of test of details risk, and a decrease in audit assurance needed from tests of details.

2 Reliance on the Internal Control Structure

Information about the effectiveness of the internal control structure at preventing or detecting material misstatements is a second source of information that the auditor evaluates during the course of an audit. When the auditor assesses control risk at less than 100%, he/she is relying, to some extent, on the internal control structure to prevent or detect material misstatements in the accounting records. The more effective the structure, the lower the auditor's assessed level of control risk. As control risk decreases, the amount of substantive

nce needed to achieve the desired level of audit assurance
ases.

The second standard of field work recognizes that the extent of substantive tests required to obtain sufficient evidential matter under the third standard should vary inversely with the auditor's reliance on internal accounting control. (AICPA [1986], AU 350.19)

The internal accounting control structure consists of three
ments: (1) the control environment, (2) the accounting system, and
control procedures. The control environment consists of
ment's control methods, management's philosophy and operating
the organizational structure, personnel policies, and if present,
ternal audit function and the audit committee (Arens & Loebbecke,
). The control environment reflects management's general attitude
philosophy about the importance of controls.

The accounting system consists of:

.. the methods and records established to identify,
assemble, analyze, classify, record, and report an entity's
ransactions and to maintain accountability for the related
ssets and liabilities. (AICPA, [1988], SAS 55).

ounting system should be designed and executed in a way that
the entity to identify, classify, record, report, and analyze
ctions correctly and to account for the entity's assets. An
ve accounting system, in an internal control sense, is one which
that all recorded transactions are valid, authorized, and
valued and classified, and that all valid transactions are
at the proper time and properly posted and summarized in the
ng records.

ontrol procedures consist of the policies and procedures
ned by the entity to ensure that the objectives of management

Examples of such procedures include segregation of duties, internal control over assets and accounting records, internal evaluation of performance, and procedures to ensure proper authorization of transactions.

The auditor is required, on every audit, to obtain an understanding of each of the three components of the client's internal structure. The auditor begins by gathering evidence to help evaluate whether the nature and design of the client's controls and procedures are conducive to the prevention and detection of errors, assuming the policies and procedures have been operating as intended throughout the period under audit. Based on this understanding, the auditor makes an initial assessment of control risk. If the auditor believes the design of the control structure to be ineffective at preventing or detecting misstatements, control risk would be assessed at the maximum level and no tests of controls would be performed. On the other hand, if the auditor believes the control structure may be effective at preventing and detecting errors, the auditor then needs to gather additional evidence, through tests of controls, that the controls have been operating effectively throughout the period. If tests of controls indicate the control structure has been effective, then the auditor would assess control risk at some level less than 100%, and the amount of substantive audit evidence necessary to achieve the desired level of audit risk could be reduced.

As greater the reliance on internal accounting control... , the greater the allowable risk of incorrect acceptance for the substantive test of details being planned and, thus, the smaller the required sample size for the substantive test of details. For example, if the auditor relies neither on internal accounting control nor on other substantive tests directed toward the same specific audit objective, he should

allow for a low risk of incorrect acceptance for the substantive test of details. Thus, the auditor would select a larger sample for the test of details than if he allowed for a higher risk of incorrect acceptance. (AICPA [1986], AU 350.19).

Reliance on Analytical Procedures

Analytical procedures are defined in Statement on Auditing Standards (SAS) #56 (AICPA [1988]) as "evaluations of financial information made by a study of plausible relationships among financial and nonfinancial data ... involving comparisons of recorded amounts to expectations developed by the auditor." If an actual recorded amount differs significantly from the auditor's expected value, then the auditor's perceived likelihood of a financial statement error increases. The importance of analytical procedures is emphasized by SAS #56 which requires their use in the planning and final review stages of every audit.

When used in the early planning stages of an audit, analytical procedures aid the auditor in deciding the nature, extent, and timing of auditing procedures by alerting the auditor to areas that may require special attention. During the course of the audit, analytical procedures may be used as substantive audit evidence to provide audit evidence relating to a particular audit objective. In the final review of an audit, analytical procedures provide an overall review of financial information and aid in evaluating the appropriateness of audit conclusions reached.

If evidence is obtained through analytical procedures that suggests there is a high likelihood that a particular account is materially misstated, the auditor would expand his/her tests of details on that account. On the other hand, if the results of analytical procedures suggest a low likelihood of misstatement, this would provide

auditor with a degree of assurance that the account balance is free from material misstatement. As the auditor's reliance on analytical procedures to detect material misstatements in the accounting records increases, the extent of other audit testing may be reduced.

Business Risk, Acceptable Audit Risk, and the External Environment

The concept of business risk has been used in the auditing literature in two contexts -- the client's business risk and the auditor's business risk. In the first context, business risk is the probability that the client will become insolvent. In the second context, business risk is the risk that an auditor will suffer a loss to her professional practice or reputation because of a relationship with a client which has become insolvent. Such losses might result from things as litigation, negative publicity, or professional sanctions. This dissertation will use the term business risk in this context.

Brumfield, Elliott and Jacobson [1983] list several factors which increase business risk (both the client's and the auditor's business risk) including: the economy in which the company operates, the industry in which the company operates, the company's financial position, operating performance, the company's existing or potential litigation, and ownership of the company. The level of business risk increases as the economy in which the client operates becomes depressed, as the client's industry becomes more volatile. Similarly, business risk increases as the client's financial position weakens, or when there is significant pending litigation. This increase in business risk is in part, to the increased risk that the client may experience

financial trouble in the near future. When firms experience financial trouble, owners and creditors may seek to recoup some of their losses through the judicial system. This, in turn, increases the likelihood the auditor may be required to defend the adequacy of the audit in court. Similarly, when the client is publicly owned, business risk is generally greater than with non-public clients. This is because the greater the number of external users of the financial statements, ceteris paribus, the greater the auditor's business risk.

Audit risk is closely related to business risk in that the auditor's perceived level of business risk may influence the auditor's level of acceptable audit risk. Brumfield, et. al. [1983] describe how business risk may influence acceptable audit risk:

Although the auditor can't change the components of audit risk, audit risk itself may be set lower (more stringently) in response to business risk. In other words, the auditor may want additional protection against the risk of issuing an incorrect opinion because of perceived high levels of business risk. Setting audit risk more stringently in this way will ultimately affect detection risk.

If the auditor believes there is a high level of business risk, the auditor may seek additional protection against the risk of issuing an unqualified opinion on materially misstated financial statements by lowering the acceptable level of audit risk. As the auditor's acceptable level of audit risk decreases, he/she will need to increase the extent of audit testing in order to increase the likelihood of detecting all material misstatements in the financial statements. The current research examines the roles of each of the four types of information discussed above - information about the internal control environment, internal control structure, analytical procedures, and external environmental factors affecting business risk -

auditors' audit planning judgments. As part of this examination, the relative importance of each of these types of information on audit planning judgments is determined, their interactive effects on these judgments is investigated, and an analysis of the effects of professional experience on audit planning judgments is performed.

The remainder of this chapter is organized as follows. The motivation for the research is presented in the next section. This is followed by discussions of the research scope and objectives. The chapter concludes with an outline of the balance of the dissertation.

MOTIVATION OF THE RESEARCH

Previous studies of auditors' program planning judgments have attempted to identify the way in which certain factors influence those judgments (e.g., Joyce [1976], Gaumnitz, Nunamaker, Surdick, & Thomas, Kaplan & Reckers [1984], Kaplan [1985], and Srinidhi & Elyi [1986]). The main focus of these studies has generally been the effects of components of the internal control structure on planning judgments.

These studies have increased our understanding of the judgment processes involved in audit program planning, and have identified some factors auditors consider important in audit planning. The present study intends to extend our understanding of auditors' judgment processes involved in the assessment of the appropriate extent of audit

To do this, the scope of the study is expanded to include external factors (in addition to internal controls), and the impact of professional experience on these judgments is investigated.

2.1 Restrictions in Scope of Previous Research on Audit Program Planning

Prior research on auditors' judgment processes has tended to focus on the relative weighting of informational cues and on the level of inter-auditor judgmental agreement. Many of these studies have examined the effects of individual controls on auditors' evaluations of the internal control structure. In studies of the judgment processes involved in audit program planning, the attention has continued to be on how individual elements of internal controls affect program planning judgments. Given the volume of this literature, examining the effects of individual internal controls on auditors' evaluations of internal control structures or on audit planning judgments no longer appears to offer much additional benefit or interest. While internal control factors may play an important role in audit program planning, it seems unreasonable to believe that auditors take into account other factors as well. However, little attention has been given to studying the effects of factors other than internal controls.

The accounting literature, while stopping short of detailing all matters auditors should consider when planning an audit program, has identified several factors which are deemed relevant to audit planning. One such factor is the client's "operating characteristics" affecting the likelihood of the occurrence of material misstatements in the entity's accounting records (AICPA [1986]). Other factors which auditors should consider include (1) matters relating to the entity's business and the industry in which it operates, (2) analytical procedures such as analyses of financial ratios and trends, and (3) internal accounting control procedures (Ibid [1986]). It should be

ed that a recent addition to the current authoritative auditing standards now requires the use of analytical procedures in the planning of all audits (AICPA [1988]). It is possible that the exclusion of factors such as those above which relate to the client's operating environment and financial condition may have produced artificiality in the judgment environment of prior studies. In order to provide subjects of this study with a richer judgment environment than has typically been employed in auditing judgment studies in the past, the four specific factors mentioned above will be included as variables. The experimental design to be used will allow us not only to examine the individual effects of each of these variables, but also to assess their combined (interactive) effects. Such an analysis is not possible when only one factor is studied at a time.

A second issue relevant to audit program planning which has received limited attention is the manner in which auditors integrate multiple cues when forming a single judgment. What little attention has been given to this issue has generally taken the form of using analysis of variance to test for significant interactions among individual control factors. Overall, the results of previous research tend to indicate that auditors exhibit very little, if any, configurality in cue utilization patterns. However, the lack of significance in the tests of interactions still leaves some ambiguity in the nature of auditors' evidence integration functions. While the extant auditing literature appears to assume that a nonsignificant interaction implies that auditors integrate evidence via a summation process, there is another possibility which has not been investigated. It is possible that auditors employ an averaging process, rather than

integrating evidence through a summation process. As is discussed in Chapter Three, this distinction is not trivial and warrants investigation. Accordingly, this study attempts to extend our understanding of the way in which auditors integrate evidence by testing two processes of evidence integration: a summing process and an averaging process.

SCOPE OF THE RESEARCH

1.1 Stages in the Audit Planning Judgment

In order to facilitate the study of auditors' judgment processes, this study conceptualizes auditors' judgments of the appropriate extent of testing of balances as consisting of three unique stages (adopted in part from Srinidhi & Vasarhelyi [1986]):

Identification & Evaluation Stage.

This stage consists of the identification and evaluation of the components of the factors considered relevant to the planning of tests of details of balances (e.g., if one factor considered relevant to the planning of tests of details of balances is the effectiveness of the internal control structure, then the components of that factor would be the individual internal controls within that structure. These components would be evaluated via tests of controls if considered appropriate).

Components Integration Stage.

This stage consists of the weighting and integration of the components of each factor considered relevant to the planning of tests of balances into a judgment of the state of that factor (e.g., the results of the individual tests of the internal

controls would be weighted and integrated into a judgment concerning the effectiveness of the internal control structure. Similarly, the components of other factors would be integrated into single judgments concerning the states of those factors).

Factors Integration Stage.

This stage consists of the weighting and integration of the judgments of the state of each relevant factor into a judgment of the appropriate extent of tests of details of balances (e.g., the evaluation of the internal control structure, the desired level of audit assurance, and other relevant factors would be integrated into a single judgment of the appropriate extent of testing).

Figure 1.1 illustrates the above discussion of the manner in which audit evidence is integrated leading up to the audit planning judgment concerning the appropriate extent of tests of details to perform.

This research concentrates on the latter two stages: the elements integration stage and the factors integration stage. In the elements integration stage, the focus is on the manner in which auditors integrate information about individual internal controls when judging the effectiveness of the internal control structure. Given the failure of most previous auditing judgment studies to find significant configurality in the judgment processes of auditors, the conclusion is that auditors will use an additive integration process in the course of the study. However, the author is aware of no prior judgment studies which have investigated whether auditors use averaging or a summing process when integrating audit evidence. Nor is the author aware of any theory which makes explicit predictions concerning averaging- vs. summing-type integration.

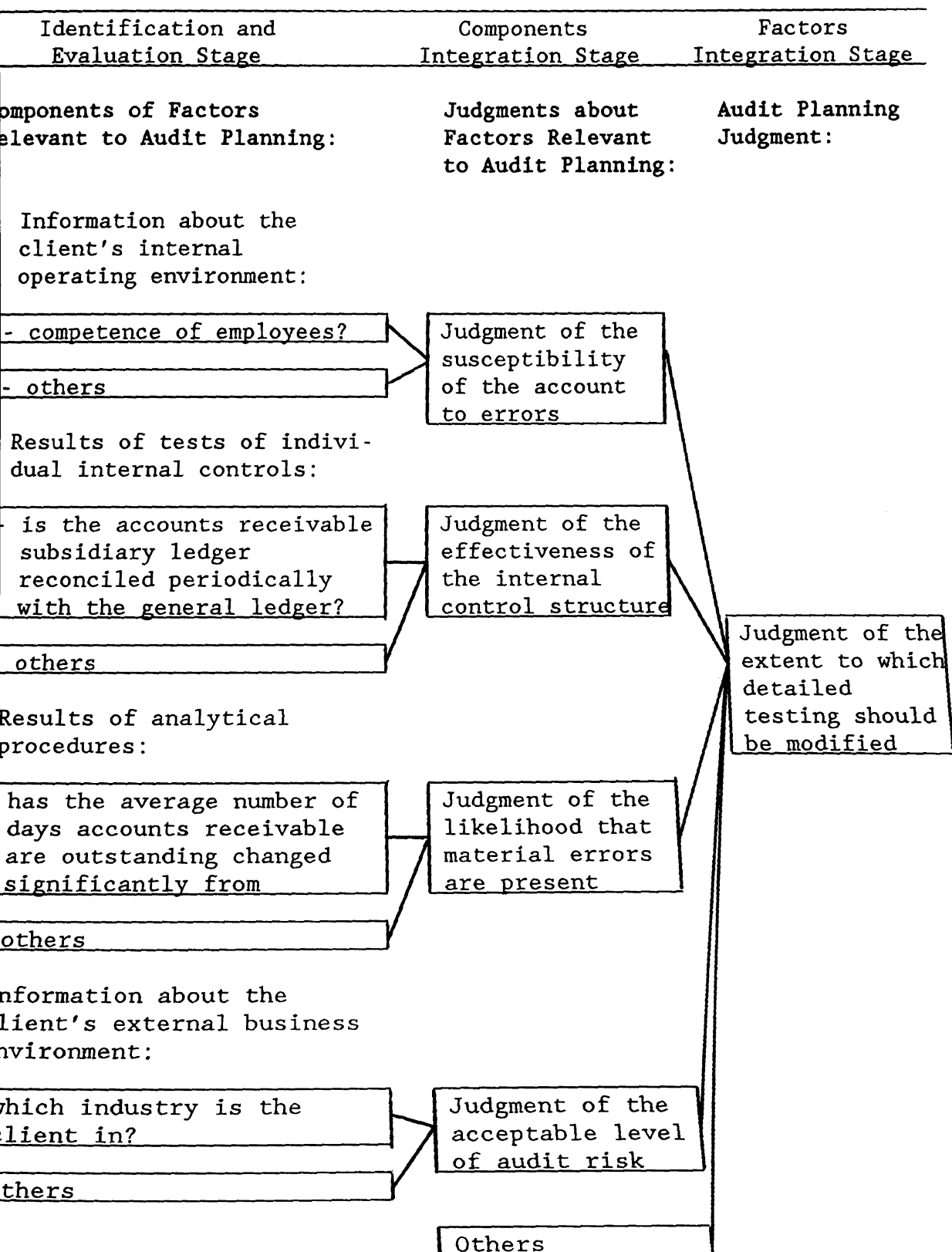


Figure 1.1
Stages of Extent-of-Tests-of-Details Judgment

Evidence from the psychology literature, specifically research based on information integration theory in the areas of person-impression formation (Anderson [1965], [1967], [1981], [1982]; Anderson and Alexander [1971]; Anderson and Birnbaum [1976]; Anderson, Lindner, and Lopes [1973]; Birnbaum [1973]; Lampel and Anderson (1968); Himmelfarb [1973]) and attitude change (Anderson [1973]; Anderson & Graesser [1976]; Sawyers and Anderson [1971]) has tended to favor an averaging rule. However, the extent of generalizability of the results of these studies to auditing contexts is currently unknown.

In the factors integration stage, the focus is on the way auditors use and integrate information concerning (1) conditions in the client's internal operating environment, (2) the effectiveness of the internal control structure, (3) the results of analytical procedures, and (4) the client's external business environment, when forming a single judgment regarding the extent of tests of details of balances to perform. Hypothesized relationships between these factors and auditors' judgments and the appropriate extent of audit testing are developed in Chapter Two.

OBJECTIVES OF THE RESEARCH

The overall objective of this research is to increase our understanding of (1) the manner in which auditors use and integrate information in audit planning and in internal control structure evaluation and (2) the role of audit experience in audit planning and in internal control structure evaluation. Specifically, this research defines the following five issues:

) the manner in which the internal operating environment, the auditor's assessment of the internal control structure, the results of analytical procedures, and the external business environment affect auditors' judgments of the appropriate extent of detailed testing of accounts receivable;

) the relative weighting, or importance, of each of the above items in auditor's extent-of-detailed-testing judgments;

) the interactive effects of the internal operating environment, the internal control structure, analytical procedures, and the external environment on extent-of-detailed-testing judgments;

) the impact of professional experience on auditors' program planning judgments, including a study of the relationship between professional experience and judgmental agreement between auditors, analysis of potential sources of disagreements between auditors, and how and where professional expertise relating to audit program planning is developed; and

) the manner in which auditors integrate tests of individual internal controls when evaluating the effectiveness of the internal control structure.

results of this research are intended to increase our understanding of audit judgment, and the effects of audit experience on audit judgment, in two areas -- the assessment of the appropriate extent of tests of details of balances, and the evaluation of the effectiveness of internal control structure.

ORGANIZATION OF THE DISSERTATION

The remainder of the dissertation consists of four chapters. Chapter Two consists of a description and discussion of previous research and the professional literature in the areas of audit planning and the effects of professional experience on auditors' judgments. Research hypotheses are developed as part of this discussion. Chapter Three consists of a detailed description of an experiment performed to investigate the research hypotheses. Chapter Four, which consists of a presentation and discussion of the results of the experiment, is divided into three sections. The use and integration of information relating to the internal operating environment, the internal control structure, analytical procedures, and the external environment in audit planning is described in Section One. The effects of professional experience on audit-planning judgments is discussed in Section Two. The manner in which tests of individual internal controls are integrated when evaluating the control structure is described in Section Three. Chapter Five consists of a summary of the findings of the research, a discussion of the limitations of the research, and suggestions for future research.

CHAPTER 2 - REVIEW OF RELATED RESEARCH AND DEVELOPMENT OF HYPOTHESES

In audit planning, a key judgment which needs to be made pertains to the amount of assurance the auditor feels is necessary that the financial statements are free from material errors before an unqualified opinion can be issued. This desired level of assurance can be thought of as the complement of the auditor's acceptable level of audit risk. The auditor relies on various factors to reduce the level of risk that material misstatements have occurred in the client's accounting records and have remained undetected. Factors the auditor may choose to rely on, if warranted, include the client's internal control structure (to reduce control risk), and the client's internal operating environment, other than internal controls (to reduce inherent risk). These two factors influence the risk that material errors will occur in the accounting process and go undetected by the client. This risk is not controllable by the auditor. To further reduce the level of audit risk, the auditor relies on analytical procedures and tests of details to detect errors which may have occurred in the accounting process and which have not been detected by the internal control structure. These two factors influence the risk that the auditor will fail to detect errors which have occurred. This risk is controllable by the auditor. Once the levels of inherent and control risk have been assessed by the auditor, he/she will perform analytical procedures or tests of details, or both, until his/her acceptable level of audit risk has been achieved. The primary focus of the current research is on the manner in which the clients' internal operating environment, the internal control structure,

results of analytical procedures, the external business environment, and professional audit experience affect the auditor's judgment concerning the appropriate extent of tests of details.

This chapter consists of a review of prior research related to the issues addressed by this dissertation. Integrated within this review is the development of the hypotheses of the current research. A review of previous research on the effects of (1) the client's internal operating environment, (2) the internal control structure, (3) the auditor's analytical procedures, and (4) the client's external business environment on auditors' judgments is presented in sections 2.1 through 2.4, respectively. Hypotheses concerning the effects of each of these four factors on audit planning judgments also are presented in sections 2.1 through 2.4. In sections 2.4 through 2.7, hypotheses are developed concerning the effects of interdependencies between the factors. In section 2.8, a review of the effects-of-audit-expertise literature is presented along with the current research's hypothesis concerning the effects of audit experience and task complexity on audit planning judgments.

2.1 AUDIT PLANNING AND THE INTERNAL OPERATING ENVIRONMENT

Kaplan and Reckers [1984] examined two factors which may potentially influence auditor's judgments of the likelihood of errors - management integrity and control consciousness. Thirty-one senior auditors and twenty-nine managers were provided with information about management integrity and internal controls over inventory, along with background information. The subjects then were asked to indicate the probability that the accounts receivable system would lead to a material

error. Control consciousness was operationalized by manipulating the strength of controls over inventory, rather than over receivables - the account under audit. Next, subjects were provided with a completed internal control questionnaire for receivables and asked to reassess the probability of errors in accounts receivable. This second response was elicited to examine whether the explicit information about controls over receivables would alter the effects of management integrity or control consciousness on auditors' probability assessments. A $2 \times 2 \times 2$ ANCOVA was used to analyze the subjects' responses. Management integrity, control consciousness, and subject rank were the independent factors in the design, general practice priors were the covariate, and auditors' initial and revised assessments of the probability of errors in receivables were the dependent variables. Kaplan and Reckers found that management integrity did not significantly influence auditors' judgments of the likelihood of errors. In addition, control consciousness was significant only for audit seniors' initial assessments, but not for their revised assessments. Control consciousness was not a significant factor in managers' judgments. Kaplan and Reckers attributed their nonsignificant results to large decision variances across auditors and recommended the use of dependent measures with less response variability in future research. Their suggestion is implemented in this dissertation.

Kaplan and Reckers' findings are not consistent with audit guidelines provided in the professional literature. Current authoritative auditing standards state that, when planning the audit, auditors should consider factors that affect the likelihood of the

occurrence of errors or irregularities in the accounting records. One such factor is the integrity of management.

The independent auditor's plan for an examination in accordance with generally accepted auditing standards is influenced by the possibility of material errors or irregularities. ... The scope of the auditor's examination would be affected by his consideration of internal accounting control, by the results of his substantive tests, and by circumstances that raise questions concerning *the integrity of management*. (emphasis added) (AICPA [1986] AU 327.06).

Kaplan and Reckers' study did not find evidence that auditors' judgments are influenced by the integrity of management. However, one can not conclude from their results that auditors' judgments are *not* influenced by management integrity. Failure to reject the null hypothesis of no effect does not indicate the alternative hypothesis is false. Therefore, judgment must be reserved as to whether the integrity of management influences auditors' assessments of the likelihood of errors.

Libby, Artman, and Willingham (LAW) [1985] also examined the influence on auditors' judgments of factors which affect the inherent likelihood of errors. In particular, LAW studied the interactive effects of the inherent likelihood of errors with (1) internal control strength and (2) compliance test strength on auditors' willingness to rely on the system of internal controls. To examine these effects, LAW manipulated the strength of controls and the strength of compliance tests over two accounting processes with differing levels of susceptibility to errors. The accounting processes consisted of a manual process for preparing purchase input forms from a voucher package and an automated process in which a computer program updates the accounts payable file and produces the weekly voucher report. LAW found

that differences in process susceptibility significantly influence the effects of control strength and compliance test strength on auditors' control reliance ratings.

Although LAW provide interesting evidence of interactive effects between the inherent likelihood of errors and control strength on control reliance ratings, their study does not examine the separate or interactive effects on auditors' extent-of-testing judgments. The current research extends the LAW study by examining these effects¹, as well as the interactive effects with analytical procedures and the client's external business environment.

The professional auditing literature suggests that conditions which affect the likelihood of material errors or irregularities should influence the auditor's judgment of the appropriate extent of substantive testing.

In planning the examination, the auditor should consider, among other matters: ... conditions that may require extension or modification of audit tests, such as the possibility of material errors or irregularities or the existence of related party transactions. (AICPA [1986] AU 311.03).

In addition, the audit approach of at least one of the Big-eight firms includes a consideration of factors in the operating environment when planning the extent of substantive testing. Grobstein and Craig [1984], in their description of the Ernst & Whinney risk analysis approach to auditing, state that the first step in the firm's Specific Risk Analysis is to identify specific factors in the operating environment which "may influence the assessment of the inherent risk associated with specific accounts or transactions." These environmental factors, along with materiality considerations, analytical procedures,

and an evaluation of the control structure then provide the basis for determining the extent of audit procedures.

Gibbins and Wolf [1982] have identified several environmental factors which auditors perceive as being important factors to consider prior to the beginning of an audit. Using a questionnaire approach, Gibbins and Wolf asked auditors to indicate "which five items from the list of potential components would be the best predictors of whether significant problems are likely to occur in conducting this year's audit." Factors which auditors generally indicated were most important included: plans for sale or major financing of client company, availability/continuity of audit personnel, the client's accounting staff, the client's accounting system, the client's top management, the client's profitability, the client's internal control, materiality considerations, and the client's financial position. The current research extends the Gibbins and Wolf [1982] study by examining the aggregate effects of several of the above factors on auditors' audit planning judgments. In the experiment performed in the current research, information about each of these factors was either provided to subjects in the instrument's background information, or was used to operationalize the independent variables. For instance, in the operationalization of the internal operating environment, subjects were presented with information regarding the competence and turn-over rate of client personnel, and various characteristics of client management, along with other information (see Appendix C).

Based on the preceding analysis, the expected effect of the client's internal operating environment on auditors' extent of testing judgments is as follows. If conditions in the internal operating

environment change from year to year or from client to client in a way which increases the likelihood of errors, the reliance the auditor is willing to place on the internal operating environment to prevent the occurrence of material misstatements in the account being audited decreases, or may even be nonexistent. As the susceptibility of an account to material error increases, inherent risk increases. Unless internal controls are 100% effective at preventing or detecting misstatements, one could not expect the client's internal control structure to completely counterbalance the increase in account susceptibility. Thus, the auditor becomes less confident that the financial statements are fairly stated. To achieve the acceptable level of audit risk, the auditor needs to increase the amount of audit evidence he/she gathers. Given that tests of details of balances are a primary source of audit evidence and are the most persuasive form of evidence, the planned extent of tests of details of balances is expected to increase in at least partial response to the need for additional audit evidence. Accordingly, a direct relationship between the susceptibility of an account balance to material error and the amount of planned audit testing is predicted. As the susceptibility of an account balance to material errors increases, the amount of planned audit evidence relating to that account's balance is expected to increase also. The analysis above leads to the first hypothesis:

- H₁: The planned extent of tests of details of balances will be *directly* related to conditions in the client's internal operating environment which increase the likelihood of the occurrence of material misstatements.

AUDIT PLANNING AND THE INTERNAL CONTROL STRUCTURE

When planning the audit, the auditor should consider his/her anticipated reliance on internal accounting controls. Several prior studies have investigated the effects of individual internal accounting controls on auditors' evaluations of the internal control structure (Ashton [1974], Ashton & Brown [1980], Ashton & Kramer [1980], Hamilton & Wright [1982], and Nanni [1984]). Other researchers have studied the effects of internal control attributes on sample size judgments (Ali & Moff [1971], and Mock & Turner [1981]), and the effects of internal control characteristics on audit program planning (Joyce [1976], Kohnitz, Nunamaker, Surdick, & Thomas [1982], Kaplan [1985], and Nidhi & Vasarhelyi [1986]). The focus of these studies has generally been on: (1) the consistency of judgment, both in terms of the consistency of the individual auditor's judgments and judgment consensus across auditors, and (2) the relative importance, or weight, of individual controls in the auditors' judgments of interest. The current study, on the other hand, will treat the effectiveness of controls as an aggregate variable. The interest in this portion of the study is on how the effectiveness of the relevant substructure of internal controls affects the auditor's judgment of the appropriate extent of tests of details of balances; not on how specific individual internal controls influence auditors' evaluations of the internal control structure.

There has been only one study which has examined the effects of internal-control effectiveness, as an aggregate variable, on audit planning decisions. Kaplan [1985] performed an experiment in which 84 auditors were presented with background information about an audit client and three internal control scenarios. Information about the

client's environment was varied in terms of stability and was treated as a between-subjects variable. Half of the subjects were asked to rate the strength of the internal control structure on a scale of one to seven and to indicate the number of audit hours they would plan for accounts receivable. The remaining subjects were simply asked to indicate their planned audit hours. They were not explicitly asked to evaluate internal control strength. Kaplan hypothesized that requiring subjects to evaluate the strength of the internal control system prior to planning the number of audit hours would affect planned audit hours and also affect the variance across auditors. Neither of these hypotheses was supported by the results of the experiment. Kaplan also hypothesized that the strength of the internal control system would influence planned audit hours. This hypothesis was supported. However, a significant internal-control-strength x environment interaction was found, indicating that the extent to which the strength of the internal control system influences planned audit hours depends, in part, on the stability of the client's environment. Like the Kaplan study, the current research examines the effects of the internal control structure on planned audit hours, and extends the Kaplan study by investigating potential interactive effects between the internal control structure, the results of analytical procedures, and factors affecting the inherent likelihood of misstatements.

Based on the preceding analysis, the expected effect of the effectiveness of the internal control structure on auditors' extent of testing judgments is as follows. As the effectiveness of an internal control structure improves, the likelihood that the internal controls will prevent errors from occurring and detect material errors which may

have occurred increases. Accordingly, the auditor, after testing the effectiveness of controls, may rely on the internal control structure to reduce control risk, which in turn reduces audit risk, *ceteris paribus*. On the other hand, if the internal control structure is not operating as planned or is poorly designed, its effectiveness at preventing and detecting errors in the accounting system is diminished. Thus the auditor may be less confident that the financial statements are error-free. As the reliance the auditor is able to place on the control structure to prevent or detect errors decreases, the extent of substantive testing needs to increase to achieve the acceptable level of audit risk. Therefore, an inverse relationship is expected between the effectiveness of the internal control structure and the extent of tests of details of balances. This expectation is supported by the following statement in the professional auditing literature:

The greater the reliance on internal accounting control... , the greater the allowable risk of incorrect acceptance for the substantive test of details being planned and, thus, the smaller the required sample size for the substantive test of details. (AICPA [1986], AU 350.19).

The analysis above leads to the second hypothesis:

H₂: The planned extent of tests of details of balances will be *inversely* related to the effectiveness of the internal control structure.

2.3 AUDIT PLANNING AND ANALYTICAL PROCEDURES

In addition to considering the client's internal operating environment and the internal control structure, the auditor should, when planning the audit, take into consideration the client's financial trends and ratios (AICPA [1986] AU 311.07). Analytical procedures, such as trend and ratio analysis, are becoming increasingly important in

auditing as a cost-effective method of identifying areas where misstatements are likely to be present and their use is expected to increase in the future (Biggs and Wild [1984]; Tabor and Willis [1985]). The usefulness of analytical procedures is apparent in Hylas and Ashton's [1982] study of how auditors discovered various errors in the financial records of actual clients. Hylas and Ashton reported that analytical procedures, in combination with discussions with client personnel and the auditor's expectations based on prior years' audits, signalled 45% of the 281 errors examined in their study. This importance of analytical procedures has been reemphasized by Statement on Auditing Standards #56 (AICPA [1988]) which now requires the use of analytical procedures in the planning and final review stages of every audit.

Analytical procedures have been the focus of much attention in the accounting literature. The majority of this attention has centered around researchers' efforts to devise and test better methods of conducting analytical procedures (Dugan, Gentry, & Shriver [1985]; Kinney [1978], [1979], [1987]; Kinney & Salamon [1982]; Kinney, Salamon, & Uecker [1986]; Lev [1980]; Loebbecke & Steinbart [1987]; and Stringer [1975]). The objective of this line of research has generally been to design statistical, usually regression-based, models for detecting unusual fluctuations in account balances, ratios, and trends, and to test the effectiveness of various analytical procedures at detecting material misstatements in account balances. A second line of research has focused on identifying the types of analytical procedures selected and applied by practitioners (Arrington, Hillison, & Jensen [1984]; Biggs and Wild [1984]; Daroca and Holder [1985]; Holder [1983]; and

Tabor and Willis [1985]). Although both areas of research are important and have made significant contributions to the literature, they provide no insight into the judgment processes employed by auditors involved in the performance of analytical procedures. While professional judgment is the key component in virtually all areas of auditing, including analytical procedures (Biggs and Wild [1984]), there have been only a few studies which have focused on audit judgment in an analytical procedures context. These studies are discussed below.

Kinney and Uecker [1982] were the first researchers to examine the judgment process behind the performance of analytical procedures. Their study focused on the anchoring and adjustment bias in auditors' judgments. One hundred seventy nine senior auditors in Big-8 accounting firms were presented with audited values for sales, cost of goods sold, and gross profit for the previous two audit years. Subjects were also presented with the current year's unaudited values and were asked to indicate a range of values for the gross profit percentage beyond which they believed an investigation of the unaudited values should be conducted. Consistent with the authors' hypotheses and with research in other judgment settings in the psychology literature (Tversky and Kahneman [1974]), auditors were found to exhibit the anchoring and adjustment bias in this analytical-procedures task. Specifically, auditors' "noninvestigation regions" for the gross profit percentage were inappropriately influenced by the auditor's knowledge of the current year's unaudited gross profit percentage. The auditors' noninvestigation regions tended to be centered on the unaudited book values. The results of Kinney & Uecker's study have important implications for practitioners since this type of judgmental bias could

lead to an undesirable increase in the likelihood of incorrect acceptance of clients' book values (type II errors). However, Kinney & Uecker did not examine how the outcome of analytical procedures effect the remainder of the audit. The current research examines this issue by investigating the effects of the results of analytical procedures, both alone and in combination with the internal control structure and the internal and external environments, on auditors' judgments of the appropriate extent of detailed testing.

Biggs and Wild [1985] also focused on the anchoring and adjustment bias, and extended the Kinney and Uecker study in a number of ways. First, the effects of receiving more than two years of prior audited data were examined by varying the amount of data presented to subjects. The amount of data was varied across two levels: two years and five years of data. Subjects who received five years of audited data exhibited biased responses in the direction of the unaudited book values, as did subjects who received only two years of audited data. However, the magnitude of the bias was smaller for subjects who received five years of data. Second, Biggs and Wild employed a control group which allowed them to determine unambiguously which of the unaudited book values had the greater effect -- the low value or the high value. Since the high value was inconsistent with the trend in the previous years, Biggs and Wild hypothesized that the high value would have the greater effect. Their results supported this hypothesis, suggesting that when the current year's unaudited book value is inconsistent with previous years' values, the auditor's knowledge of this unaudited value will bias his/her response to a greater extent than when the unaudited value is consistent with previous years' values. This increases the

risk of not finding a material error, since the auditor is less likely to investigate the account balance. Third, Biggs and Wild examined auditors' abilities to develop accurate expected values for financial statement items in an analytical procedures setting. Subjects were presented with seven years of audited values for revenue and were asked to determine an expected value for the current year. Previous years' values followed one of six patterns: either increasing or decreasing exponential, logarithmic, or linear trends. Subjects' extrapolations for all three increasing patterns were found to be significantly underestimated, while subjects' extrapolations for all three decreasing patterns were significantly overstated. In addition, extrapolations for the exponential patterns were significantly less accurate than for either the logarithmic or linear patterns. This suggests that when account values are changing rapidly from year to year, an auditor's attempt to judgmentally extrapolate an expected value for the current year may result in systematically understated (or overstated) expected values. As did the Kinney & Uecker study, the Biggs & Wild study provides interesting information about potential biases in auditors' analytical procedures judgments. However, also like Kinney & Uecker, Biggs & Wild did not investigate the manner in which the outcome of analytical procedures affect the remainder of the audit. The current research investigates this issue.

Shields, Solomon, and Waller (SSW) [1985] also investigated the issue of anchoring and adjustment in auditors' performance of analytical procedures. Their study was motivated by the proposal that, in some cases, the auditor's use of current unaudited book values when performing analytical procedures might increase audit effectiveness. To

investigate this, SSW performed an experiment in which 62 auditors were presented with real-world auditee data and asked to provide probabilities for various account-value intervals for six account balances. Two groups of subjects were used: an experimental group which received unaudited book values that were either fairly stated or materially misstated and a control group which received no unaudited book values. Results indicated that subjects in the experimental group relied on the unaudited book values regardless of whether they were correct or misstated. Subjects in the experimental group produced account-value judgments which were more accurate than subjects' judgments in the control group when the unaudited values were fairly stated and less accurate judgments when the unaudited values were misstated. These results suggest that when the unaudited book values are fairly stated, the auditor's knowledge and use of unaudited book values may result in a more efficient audit; i.e., less incorrect rejections of account balances. However, when the unaudited account balance is misstated, the auditor's use of that balance may result in a less effective audit as the likelihood that the account balance will be incorrectly accepted increases. An examination of the effect of the outcome of the analytical procedures on the rest of the audit was beyond the scope of SSW's study. The current research extends their study by examining the manner in which the results of analytical procedures effect the performance of the audit.

Two studies of auditors' judgments in analytical procedures settings which do not focus on auditors' anchoring and adjustment behavior are by Blocher, Esposito, and Willingham [1983] and by Libby [1985]. Blocher, Esposito, and Willingham [1983] performed an

experiment which examined the effects of the prior year's audit program and a checklist of suggested analytical procedures on auditors' judgments of the appropriate extent of analytical procedures in the payroll area. Results of the experiment indicated that there was much variability across auditors' judgments of the number of hours to be budgeted for the performance of analytical procedures. However, neither the prior year's audit program nor the checklist of suggested analytical procedures significantly affected auditors' judgments regarding budgeted hours for analytical procedures. As in the previously discussed studies, Blocher, et. al. did not examine how the outcome of analytical procedures affect the performance of the audit. The current study addresses this issue.

Libby [1985] examined how auditors develop initial hypotheses, or explanations of unusual fluctuations, in analytical procedures settings. His primary emphasis was on studying the roles played by auditors' task-related knowledge and memory of possible causes of financial statement errors. Libby found that the likelihood that a particular financial statement error will be offered by the auditor as an explanation of an unusual fluctuation is influenced by the auditor's perceived frequency of that error. Related to this was the finding that auditors perceive that errors which overstate income occur more frequently than errors which understate income. Libby's study is alone in its focus on the roles of task-specific knowledge and memory in an analytical procedures setting, and offers unique insight into the importance of professional experience in analyzing the results of analytical procedures.

While the preceding studies of auditor judgment in analytical-procedures settings have provided interesting contributions to the

literature and have increased our understanding of auditors' judgment processes, they have ignored the effects of the outcome of analytical procedures on the remainder of the audit. The current research examines this issue by investigating the effects of the results of analytical procedures, both alone and in combination with the internal control structure and the internal and external environments, on auditors' judgments of the appropriate extent of detailed testing.

Analytical procedures can be used in the early stages of an audit as an "attention director" to alert the auditor to areas which may present higher risks of error. They may also be used as substantive audit evidence to provide assurance relating to a particular audit objective. In the final review stages of the audit, analytical procedures can be used as a means of evaluating the audit conclusions reached and to identify any remaining unusual relationships in the financial statements which have not been previously identified (AICPA [1988]). Regardless of the particular stage of the audit in which analytical procedures are employed, the outcome of analytical procedures may potentially affect the auditor's judgment of the appropriate extent of tests of details of balances to perform. When used in the planning stages of an audit as an attention director, unfavorable results in a given analytical procedure may result in the auditor performing expanded tests of details of the balances of the associated accounts. When used as substantive audit evidence, favorable (unfavorable) results in a given analytical procedure may result in the auditor decreasing (increasing), or, depending on the particular account involved, eliminating further substantive tests of the account. Accordingly, an inverse relationship between the favorableness of the results of

analytical procedures and the planned extent of tests of details of balances is predicted. The analysis above leads to the third hypothesis:

H₃: The planned extent of tests of details of balances will be *inversely* related to the favorableness of the results of analytical procedures.

2.4 AUDIT PLANNING AND THE CLIENT'S EXTERNAL BUSINESS ENVIRONMENT

The auditing literature suggests that the client's business environment should play an important role in audit planning. According to current authoritative auditing standards, "in planning the examination, the auditor should consider ... matters relating to the entity's business and the industry in which it operates." (AICPA [1986] AU 311.03)

Kaplan [1985] examined whether auditors consider the client's environment, including the client's industry, in audit planning. The environment was manipulated over three levels -- dynamic, slightly dynamic, and stable. He hypothesized that planned audit hours would be highest when the environment was dynamic, and lowest when it was stable. Planned audit hours were found to be consistently greater in the dynamic environment than in either the slightly dynamic or stable environments, and greater in the slightly dynamic environment than in the stable environment. However, a significant interaction between the environment and internal control was found. When the effects of the environment were examined under conditions of weak internal control, the environment was found to not significantly affect audit planning. Therefore, the hypothesized relationship between audit hours and the environment was

not supported. Accordingly, this researcher believes that judgment must be reserved, pending further research, as to whether conditions in the environment do indeed affect auditors' judgments concerning budgeted audit hours.

Although Kaplan's hypothesis was not supported, the auditing literature suggests that the client's business environment should influence audit planning decisions. External factors such as the client's business and industrial environment affect the auditor's business risk. Business risk can influence the planning of the audit through its effect on the auditor's desired level of assurance that the financial statements are free from material misstatements once the audit is completed and an unqualified opinion is issued.

Although the auditor can't change the components of audit risk, audit risk itself may be set lower (more stringently) in response to business risk. In other words, the auditor may want additional protection against the risk of issuing an incorrect opinion because of perceived high levels of business risk. Setting audit risk more stringently in this way will ultimately affect detection risk. (Brumfield, Elliott, and Jacobson [1983]).

In planning the audit, the auditor uses his/her professional judgment as to the appropriate level of audit assurance to be sought. There are many factors which may be considered by the auditor when establishing this desired level of assurance. Often these factors relate to the likelihood that, sometime in the future, the auditor may be called upon to defend the quality of the audit in court. If the auditor feels there is a relatively high likelihood that the audit may be brought under scrutiny, he/she may find it desirable to increase audit assurance to a higher level than that which would normally be sought under other circumstances. For example, if the auditor believes that external users

of the financial statements will be relying heavily on those statements, the auditor may want to protect himself/herself, as well as the users, by increasing the level of assurance that any material errors in the financial statements will be detected by the auditor. Such a situation may occur, for instance, when the client is about to go public. Another situation in which the auditor may find it desirable to increase audit assurance would be when the auditor believes there is a relatively high likelihood that the client may experience financial trouble in the near future. This situation, like the first, increases the likelihood that the auditor may be required to defend the adequacy of the audit. If an audit client subsequently goes bankrupt, for example, stockholders may try to recoup some of their losses by suing the auditor.

As circumstances in the client's external business environment become more uncertain or more volatile, the auditor is expected to become more concerned about the possibility of some situation or set of events occurring which may result in financial difficulty for the client. This, in turn, is expected to influence the auditor's desired level of assurance for that audit. As explained by Arens and Loebbecke [1988]:

If a client is forced to file for bankruptcy or even just suffers a significant loss after completion of the audit, there is a greater chance of the auditor being required to defend the quality of the audit than if the client were under no financial strain. There is a natural tendency for those who lose money in a bankruptcy or because of a stock price reversal to file suit against the auditor. (pg. 244).

Therefore, as the volatility in the client's external business environment increases, the auditor is expected to seek a higher level of assurance that all material errors in the financial statements will be discovered. As the desired level of audit assurance increases, the

auditor will need to accumulate additional audit evidence to attain the higher level of assurance. Inasmuch as tests of details of balances are a primary source of audit evidence, the auditor is expected to expand his/her tests of details in response to an increase in the volatility of the client's environment. Accordingly, a direct relationship between the degree of volatility in the client's business environment and the extent of tests of details of balances is predicted. The analysis above leads to the fourth hypothesis:

H₄: The planned extent of tests of details of balances will be *directly* related to the volatility in the client's external business environment.

The four preceding hypotheses predict that auditors' audit planning judgments will vary systematically with changes in (1) the client's internal operating environment, (2) the internal control structure, (3) the auditor's analytical procedures, and (4) the client's external business environment. Of additional interest is the determination of the relative importance, or weight, of each of these four factors to the auditors' audit planning judgments. Each factor's eta-squared index will be examined to help determine the relative importance of the factors. The eta-squared index for a given factor reflects the proportion of the variance in the auditors' judgments that is explained by that factor.

2.5 INTERNAL CONTROLS AND THE INHERENT LIKELIHOOD OF ERRORS

Also of interest is the manner in which the effects of each factor combine or interact with other factors in their influence on auditors' judgments. Libby, Artman, and Willingham [1985] found that as the strength of the internal control structure decreases, the auditor's

reliance on the internal controls decreases by a greater margin when the accounting process is more susceptible to errors than when the process is less susceptible. This suggests the presence of an interactive effect between the strength of internal controls and process susceptibility in auditor's internal control reliance decisions.

In the current research, a similar interactive effect is expected between control strength and account susceptibility on auditors' judgments of the extent of tests of details of balances. Given the results of Libby, Artman, and Willingham's study, changes in the effectiveness of the internal control structure are expected to have a greater effect on auditors' planned extent of tests of details when account susceptibility is high than when it is low. Intuitively, this expectation appears reasonable. If the account is very susceptible to misstatements, then the auditor's assessment of the internal control structure's ability to prevent or detect misstatements in that account could have a significant effect on the extent of testing of that account's balance. To illustrate, if the account is quite susceptible to misstatements, but the auditor believes that the internal control structure would probably detect any misstatements which may have occurred, then the auditor may feel fairly confident that the account is not misstated and will perform reduced tests of the account's balance. On the other hand, if the auditor believes that the controls are ineffective, then he/she may feel fairly confident that the account *is* misstated, and will perform correspondingly more tests of that account. However, if the account under audit *is not* very susceptible to misstatements - if there is a low likelihood that errors have occurred in that account's balance within the past period - an increase or

decrease in the effectiveness of internal controls would be expected to have little effect on the extent of tests of details of that account.

This analysis leads to the fifth hypothesis:

- H₅: An *increase* in the effectiveness of the internal control structure will result in a *larger decrease* in the extent of tests of details of balances when the client's internal operating environment is conducive to a *high* likelihood of misstatements than when the likelihood of misstatements is low.

Libby, Artman, and Willingham's findings are consistent with what the audit risk model would predict. It should be noted, however, that the hypothesis above runs counter to the audit risk model. The audit risk model predicts that changes in the effectiveness of the internal control structure would have their greatest effect on the extent of tests of details of balances when the account's susceptibility to errors is low, rather than high².

2.6 INTERNAL CONTROLS AND ANALYTICAL PROCEDURES

Cushing and Loebbecke [1983] argue that analytical procedures are likely to be less effective in detecting errors in accounting records when the internal control structure is weak, than when the structure is strong. They state:

A common analytical review procedure involves identifying unusual fluctuations for more detailed investigation based on trend data. For example, monthly charges to an account in the current year may be reviewed for reasonableness based on similar monthly charges for the preceding 24 or 36 months (the base period). There would appear to be an assumption that (1) the base period data is correct, and that (2) the current period data can't be fraudulently adjusted (normalized) in order to make these comparisons appear to be valid. Both of these assumptions are appropriate *only if IC (control risk) is low*. (emphasis added) (Cushing and Loebbecke [1983], pg. 29).

This suggests that the amount of assurance that auditors can obtain from

the results of analytical procedures depends, in part, on the effectiveness of the internal control structure. When controls are effective, results of analytical procedures may be more reliable than when controls are ineffective. The current research examines this issue.

Ashton [1974] found that when auditors discover a weakness in the internal control structure, their evaluation of the effectiveness of the structure drops very sharply, and the results of tests of other individual controls seem to have little effect on the evaluation of the structure. Similarly, in the current study, if weaknesses are found in the internal control structure, it is expected that the extent of tests of details of balances will rise sharply, and the effects of other evidence (i.e., analytical procedures) will be attenuated. When controls have not been operating effectively throughout the period under audit, the auditor is expected to place less assurance on the results of analytical procedures than when controls have been operating effectively. Accordingly, it is not expected that favorable analytical procedure results will persuade the auditor to cut back much on testing of details of balances when internal controls are not operating effectively. This analysis leads to the sixth hypothesis:

- H₆: A *decrease* in the favorableness of the results of analytical procedure will result in a *larger increase* in the extent of tests of details of balances when the internal control structure is *strong* than when it is weak.

2.7 ANALYTICAL PROCEDURES AND THE EXTERNAL BUSINESS ENVIRONMENT

Statement on Auditing Standards No. 56 suggests that the effects of analytical procedures on auditors' judgments of the appropriate

extent of tests of details of balances may depend on the degree of volatility in the client's business environment.

It is important for the auditor to understand the reasons that make relationships plausible because data sometimes appear to be related when they are not, which could lead the auditor to erroneous conclusions ... As higher levels of assurance are desired from analytical procedures, more predictable relationships are required to develop the expectation. Relationships in a stable environment are usually more predictable than relationships in a dynamic or unstable environment. (AICPA [1988]).

If the results of analytical procedures are favorable, it is expected that, in a relatively stable environment, the level of assurance that the auditor is willing to place on the results of those analytical procedures will be greater than if the environment were volatile. This would result from the auditor's greater confidence in his/her initial expectation, which in turn results from the greater predictability of relationships in the stable environment. This reasoning leads one to expect that, given favorable analytical procedure results, the extent of tests of details of balances would decrease by a greater margin in a stable environment than in a volatile environment. This analysis suggests the seventh hypothesis:

H₇: A *decrease* in the favorableness of the results of analytical procedures will result in a *larger increase* in the extent of tests of details of balances when other information suggests a *low* degree of volatility in the client's external business environment than when other information suggests a *high* degree of volatility in the client's external business environment.

2.8 EFFECTS OF PROFESSIONAL EXPERIENCE AND TASK COMPLEXITY ON AUDIT PLANNING

Previous behavioral auditing research has often focused on the effects of attributes of internal controls on audit judgments. The

audit judgment involved in the evaluation of the quality of internal controls has been examined by Ashton [1974], Ashton & Brown [1980], Ashton & Kramer [1980], Hamilton & Wright [1982], and Nanni [1984]. Ali & Duboff [1971] and Mock & Turner [1981] have studied the effects of internal control attributes on sample size judgments, and Joyce [1976], Gaumnitz, Nunamaker, Surdick, & Thomas [1982], Kaplan [1985], and Srinidhi & Vasarhelyi [1986] have focused on the effects of internal control characteristics on audit program planning. Since these studies have lacked a criterion by which to measure the accuracy of their subjects' judgments, they have tended to focus on the stability and consensus of the subjects' judgments.

The findings of several of these studies appear to contradict one another to some extent. Joyce [1976], in his study of audit program planning, found the mean level of consensus between auditors to be surprisingly low, .37, with values ranging from -.69 to .94. On the other hand, Ashton [1974] and Ashton & Brown [1980], focusing on auditors' evaluations of internal controls, reported average consensus levels of .70 and .67, respectively. Gaumnitz, et. al. [1980] also reported a mean consensus level between auditors (.62) which was considerably higher than that found by Joyce.

In an effort to explain the finding that auditors seem to disagree on audit program planning judgments but agree on assessments of the effectiveness of internal controls, Srinidhi & Vasarhelyi [1986] attempted to separate the effects of the internal-control-evaluation judgment from the audit-program-planning judgment. They found that when subjects were required to make judgments regarding the extent and timing of substantive tests given reliability ratings of the *components* of an

internal control system, consensus between auditors was fairly low, .58. However, when their subjects were asked to make the same judgment given a reliability rating of the internal control system as a whole (i.e., subjects were not required to integrate any information regarding the internal control components to form a judgment concerning the strength of the internal control system), the consensus level climbed to .94. The main conclusion of their study was that "auditors tend to disagree on how to aggregate the evidence from the system in order to arrive at a system reliability measure."

Other research in this area provides an explanation of Srinidhi & Vasarhelyi's findings. It has been shown that as the level of difficulty of a task increases (i.e., as the amount of expertise, or experience, required to complete a task increases), the level of consensus between subjects tends to decrease (e.g., Abdolmohammadi & Wright [1987]). This is especially true when using subjects whose experience/expertise levels are lower than what is normally required to perform the particular audit task being studied (Bonner [1988]). This suggests that one would expect auditors to exhibit a relatively low level of consensus when they are required to evaluate internal control components, form a judgment regarding the strength of the internal control system, and then decide how this judged strength of the internal control system should influence the extent of substantive testing. However, when auditors are asked to decide the extent of substantive testing *given* the strength of controls, their consensus level should be higher. This is because the latter judgment is relatively easier and has fewer sources of potential variations in judgments.

Furthermore, it should be noted that Joyce's [1976] reported

consensus level was still lower than Srinidhi & Vasarhelyi's consensus level for the case in which their subjects were given only the reliability ratings of the components of the internal control system. In Joyce's study, subjects were not only required to consider the effects of characteristics of the internal control system, but were also required to consider the impact of the results of two analytical procedures when deciding the appropriate extent of detail testing. Given that judgments regarding analytical risk in audit planning are usually performed by auditors at the manager level (Bonner [1988]), it could be argued that Joyce's task was more difficult (or similarly, one which required more expertise/experience) than Srinidhi & Vasarhelyi's task, thus explaining the lower level of consensus found by Joyce.

Many of the previous studies of auditors' judgment processes appear to have selected subjects without much concern for the amount of experience those subjects have, or for how the difficulty of the task the subjects are asked to perform may interact with the subjects' experience levels. However, it is reasonable to suppose that an audit manager with seven years of experience may employ different judgment strategies than a staff accountant with one year of experience. Likewise, the staff accountant may have different judgment strategies than an auditing student with no professional experience. Differences in their judgment strategies may include different cue-weighting strategies, different degrees of configural cue-utilization strategies, and different degrees of consistency in following these strategies. Given these possibilities, it may be important to consider and control for the experience levels of subjects in auditing judgment research. Consideration should also be given to the difficulty of the selected

judgment task and the staff level at which experience at that task is normally obtained.

Ashton & Kramer [1980] were among the first to examine how auditors with different levels of experience differ in their judgment strategies. Comparing the judgment processes of auditing students and practicing auditors involved in evaluations of internal control structures over payroll, Ashton & Kramer found that the patterns of cue utilization of the two groups of subjects were quite similar. Ashton & Kramer concluded that, for the evaluation of internal controls, students appear to be acceptable surrogates for experienced auditors. Apparently, professional experience does not significantly alter the judgment strategy involved in the evaluation of internal controls over payroll.

Using a task very similar to that used by Ashton [1973] and Ashton & Kramer [1980], Hamilton & Wright [1982] also investigated the effects of professional experience on the evaluation of internal controls. Their findings were consistent with those of Ashton & Kramer [1980] in that they found the amount of experience to be unrelated to judgment consensus, judgment stability, and cue-weighting, and that self-insight appears to increase with experience. Like Ashton & Kramer, Hamilton & Wright concluded that experience does not produce an improvement in the level of consensus between auditors. However, they caution that their results may not be generalizable to less-structured judgment contexts.

There are, however, studies which report findings which are not consistent with those discussed above. Nanni [1984] also examined the effects of auditor experience on internal control evaluation. Focusing on ratings of control reliability and consensus within groups of

auditors, Nanni found significant differences between the ratings and consensus levels of managers, supervisors, and seniors. Interestingly, Nanni's results indicate that seniors and managers were more conservative and exhibited less judgmental agreement among themselves than did supervisors. Nanni's post hoc explanation of these results attributes this to diminishing experience effects beyond some point, after which responsibilities and evaluation goals begin to dominate.

Experience effects on the judgment processes of auditors were also reported by Krogstad, Ettenson, & Shanteau [1984]. Focusing on materiality judgments of students, audit seniors, and partners, Krogstad, et. al. found that, contrary to the findings of Ashton & Kramer and Hamilton & Wright, students differed significantly from both seniors and partners in terms of cue usage, consensus levels, and stability of judgments. Krogstad, et. al.'s study reported that the students tended to focus on several cues, while the practicing auditors mainly emphasized the "effect on net income" cue, and that the students had significantly lower levels of judgment stability and judgment consensus than the seniors and partners. However, it should be noted that the responses of the seniors and the partners were quite similar to one another. No significant differences between those two groups were found in terms of judgment stability, consensus, or cue usage, indicating that most of the effects of experience relevant to the materiality judgment occur by the time an auditor reaches the senior level. This latter finding is inconsistent with the Nanni study which found significant differences between the judgments of seniors, supervisors, and managers. While the differences in the judgments of students and auditors found by Krogstad, et. al. also contradict the

results reported by Ashton & Kramer and Hamilton & Wright, the judgment task used by Krogstad et al. differed from that of Ashton & Kramer and Hamilton & Wright. As pointed out by Krogstad, et. al., the evaluation of internal controls is a fairly structured task with which most auditing students are familiar, while materiality judgments are less structured and require more subjective judgment. Therefore, the discrepancy of the results of the Krogstad, et. al. study from those of Ashton & Kramer and Hamilton & Wright may be due to the fact that the complexity of the task varied across the studies. The evaluation of internal controls is a relatively easy audit task, while materiality judgments are relatively more difficult, allowing the experienced auditors to exhibit the effects of their professional experience in that task.

The differences in the results of these studies underscores the importance of considering the relative difficulty of the judgment task chosen for study, and the staff level at which experience at the task is obtained, when studying the effects of experience on the judgment processes of auditors. Professional experience may be important for certain unstructured, complex tasks, but irrelevant for well-structured, simple judgments.

Task complexity and the stage at which experience at the task is obtained were explicitly considered by Abdolmohammadi & Wright [1987]. Using several judgment tasks which varied from structured to unstructured and groups of auditors which varied in their years of experience, Abdolmohammadi & Wright found that, consistent with their hypothesis, experience effects increase as task complexity increases. This was indicated by significant differences in the judgments of

experienced and inexperienced auditors when performing unstructured and semi-structured audit tasks. However, contrary to their hypothesis, significant differences between the two groups of subjects were also found in the structured task. This is attributed to an unfortunate choice of structured task - it turned out to be dependent on the judgments formed in the semi-structured task. Interpretation of their results is made somewhat difficult by the fact that the amount of experience needed for an auditor to be classified as "experienced" in their study varied across tasks. Therefore, the experience effects which were found could be attributed to subject differences other than task-specific knowledge derived from experience; e.g., keener analytical skills relevant to judgments in general.

Bonner [1988] also studied the effects of experience on auditor judgment. She attempted to correct for certain deficiencies present in the designs of previous studies. The key improvement in her research design was the use of two groups of subjects (experienced-novice) and two tasks (complex-easy). Subjects in both groups performed both tasks. This allowed a comparison of the size of the performance difference between the two groups of subjects when performing the easy task, with the size of the performance difference between the two groups when performing the complex task. This comparison of the performance *differences* of the two groups of subjects *across* tasks allows one to identify more confidently task-specific knowledge differences as the source of experience effects, and eliminates most competing possible explanations of experience effects. Consistent with her hypothesis, the effects of knowledge obtained through audit experience were greater when subjects performed the more complex of the two tasks.

The findings of the studies discussed above indicate the importance of carefully considering (1) auditing experience when gathering subjects, (2) task complexity when choosing an experimental judgment task, and (3) the staff-level at which auditors obtain the knowledge relevant to the chosen task. Furthermore, if one goal of the study is to be able to draw conclusions regarding the importance of task-specific knowledge obtained from audit experience, the design of the study and choice of subjects and tasks must be such that the researcher can rule out competing explanations of differences in the performance of experts and novices.

To study the effects of task-specific knowledge obtained through experience on audit planning judgments, the current research uses a design similar to that used by Bonner [1988]. That is, both experience level and task complexity are varied, with the focus on comparisons of performance differences of the groups of subjects *across* task complexity. In an attempt to gain an understanding of the nature and timing of the formation of expert judgment, this study compares the judgments of three groups of subjects (audit managers, seniors, and auditing students) across two tasks varying in complexity (internal accounting control evaluation and audit program planning). This is also intended to increase our understanding of the influence that professional experience and education have on the formation of such expertise. Finally, it is anticipated that this study will increase our insight into the appropriateness of using students as surrogates for practicing auditors in behavioral research.

It is expected that there should be a large difference in the consensus levels between the group of practitioners and the group of

students when forming judgments regarding the appropriate extent of testing of details of balances, and a small difference between the consensus levels of these two groups when evaluating internal controls. This expectation of a relatively small difference between the consensus levels of practitioners and students when evaluating internal controls is based in part on the assumption that students receive training in their auditing courses on which characteristics of internal control structures constitute strengths and weaknesses. However, it is expected that judgments of the appropriate extent of testing of details of balances will be more difficult than internal control evaluation *by a larger margin* for students than for practitioners. This expectation is based on the findings of previous research on audit planning and on internal control evaluation (e.g., Joyce [1976], Ashton [1974]) which tend to suggest that audit planning is a more difficult, less structured task than is the evaluation of internal controls. Professional experience is expected to play a more prominent role in audit planning than in the evaluation of internal controls. Therefore, there should be a larger difficulty differential between the experienced auditors and the students when assessing the appropriate extent of tests of details of balances than when evaluating internal controls. This analysis leads to the eighth hypothesis:

- H₈: There will be a *larger difference* between the level of judgmental agreement among experienced auditors and the level of agreement among students when assessing the appropriate extent of tests of details of balances than when evaluating the internal control structure.

In addition to comparing the level of judgmental agreement among students with the judgmental agreement among practitioners, comparisons are made of the judgmental agreement among students with that among

seniors, and the judgmental agreement of seniors with that of managers. This provides a finer understanding of how and where auditors develop the expertise and knowledge relevant to professional audit planning judgments. If auditors develop the expertise relevant to audit planning early in their professional careers (within the first one to two years), no significant differences would be expected between the level of judgmental agreement among managers and the level of agreement among seniors when judging the appropriate extent of tests of details of balances than when evaluating internal controls (i.e., managers would have no task-specific knowledge beyond that possessed by seniors). On the other hand, if managers possess more task-specific expertise than seniors in audit planning, then a larger difference would be expected between the level of judgmental agreement among managers and the level of agreement among seniors when assessing the appropriate extent of tests of details than when evaluating the control structure. These comparisons allow an assessment of the relative importance of formal education and professional experience in the process of developing professional judgment. An understanding of where and how professional judgment develops relevant to audit planning may help firms in the development of educational and training aids. Finally, this study provides evidence regarding the appropriateness of using students and/or seniors in studies of the audit program planning judgment.

For the reader's convenience, a summary of the research hypotheses developed in this chapter is presented in Figure 2.1.

- H₁: The planned extent of tests of details of balances will be *directly* related to conditions in the client's internal operating environment which increase the likelihood of the occurrence of material misstatements.
- H₂: The planned extent of tests of details of balances will be *inversely* related to the effectiveness of the internal control structure.
- H₃: The planned extent of tests of details of balances will be *inversely* related to the favorableness of the results of analytical procedures.
- H₄: The planned extent of tests of details of balances will be *directly* related to the volatility in the client's external business environment.
- H₅: An *increase* in the effectiveness of the internal control structure will result in a *larger decrease* in the extent of tests of details of balances when the client's internal operating environment is conducive to a *high* likelihood of misstatements than when the likelihood of misstatements is low.
- H₆: A *decrease* in the favorableness of the results of analytical procedures will result in a *larger increase* in the extent of tests of details of balances when the internal control structure is *strong* than when it is weak.
- H₇: A *decrease* in the favorableness of the results of analytical procedures will result in a *larger increase* in the extent of tests of details of balances when other information suggests a *low* degree of volatility in the client's external business environment than when other information suggests a *high* degree of volatility in the client's external business environment.
- H₈: There will be a *larger difference* between the level of judgmental agreement among experienced auditors and the level of agreement among students when assessing the appropriate extent of tests of details of balances than when evaluating the internal control structure.

Figure 2.1
Summary of Research Hypotheses

ENDNOTES

1. Although Libby, Artman, and Willingham assert the interactive effects of the inherent likelihood of errors and control strength on the extent of substantive testing should be the same as those observed for control reliance ratings, the audit risk model and the findings of the current research suggest otherwise. This issue is dealt with in detail in sections 2.3.5 and 4.1.3.
2. A numerical example, using two cases, will illustrate the way in which hypothesis five runs counter to the audit risk model:

$$\text{Audit Risk Model: } \text{ATD} = \frac{\text{AAR}}{\text{IR} \times \text{CR} \times \text{AnR}}$$

where: ATD = Acceptable Test of Details Risk
 AAR = Acceptable Audit Risk
 IR = Inherent Risk
 CR = Control Risk
 AnR = Analytical Risk

In case 1, assume the internal operating environment is unfavorable; i.e., inherent risk = 1.0. In case 2, assume the internal operating environment is more favorable; i.e., inherent risk = .60. Assume desired audit risk and analytical risk in both cases is .05 and 1.0, respectively.

Case 1:

$$\text{High Inherent Risk; Low Control Risk: } \text{ATD} = \frac{.05}{1.0 \times .3 \times 1.0} = .166.$$

$$\text{High Inherent Risk; High Control Risk: } \text{ATD} = \frac{.05}{1.0 \times .9 \times 1.0} = .055.$$

In case 1, a decrease in the strength of controls, i.e., a change in control risk from .3 to .9, will result in a decrease in acceptable test of details risk (which will result in an increase in audit testing) of .111 (.166 - .055).

In case 2, the same decrease in the strength of controls will result in a larger decrease in acceptable test of details risk (and a larger increase in audit testing) of .184 (.277 - .093):

Case 2:

$$\text{Low Inherent Risk; Low Control Risk: } \text{ATD} = \frac{.05}{.6 \times .3 \times 1.0} = .277.$$

$$\text{Low Inherent Risk; High Control Risk: } \text{ATD} = \frac{.05}{.6 \times .9 \times 1.0} = .093.$$

Thus, the audit risk model predicts that a change in the strength of controls will have a greater effect on the extent of testing when the likelihood of misstatements (inherent risk) is low, rather than high.

CHAPTER 3 - METHODOLOGY

3.1 INTRODUCTION

The primary goal of this study is to provide an in-depth examination of (1) the manner in which auditors use and integrate information in audit planning and in internal control structure evaluation and (2) the role of professional experience in audit planning and in internal control structure evaluation. Specifically, this dissertation examines:

- (1) the manner in which auditors integrate results of tests of individual internal controls into a single judgment of the effectiveness of the internal control substructure over the account being audited. This judgment occurs in the components-integration stage of the auditor's judgment process.
- (2) the manner in which auditors utilize and integrate evidence concerning: (i) the internal operating environment, other than internal controls, (ii) the effectiveness of the internal control structure, (iii) the outcome of analytical procedures, and (iv) the volatility of the client's external business environment, when judging the appropriate extent of tests of details of balances to be performed. This judgment occurs in the factors-integration stage of the auditor's judgment process.

This research also examines whether audit planning judgments of experienced practitioners tend to differ from those of less experienced auditors, and if so, in what aspects they differ. In particular, the

focus is on identifying (1) the stages of the audit planning judgment process in which differences between experienced auditors and novice auditors tend to occur, (2) the stages of an auditor's professional career in which he/she develops sufficient skill in making professional judgments relevant to audit planning decisions, and (3) whether differences in the judgment processes of experienced auditors and novice auditors stem from differences in cue weighting agreement, judgment consistency, and/or the extent of configurality in their judgment processes.

3.2 RESEARCH INSTRUMENT AND ADMINISTRATION OF THE EXPERIMENT

Practicing auditors and accounting students were used as subjects in an experiment designed to investigate the issues addressed by the current research. A contact person in a large, international public accounting firm was asked to participate in securing practitioner subjects for the study and in administering the instrument. The contact person was asked to seek the cooperation of senior auditors and audit managers. The set of instruments was delivered in person to the contact, who in turn, delivered an instrument to each subject. The maximum number of audit managers over whom the contact person had jurisdiction was twenty. All twenty of these managers received an instrument. While the contact person had jurisdiction over a larger number of audit seniors, the researcher desired equal numbers of seniors and managers. Accordingly, twenty seniors also received an instrument. Twenty accounting students enrolled in their first undergraduate auditing course also served as subjects in the experiment.

Subjects were asked to (1) read the instructions and background information, (2) evaluate 10 audit planning cases, and (3) evaluate 10 internal control cases. The instrument consisted of three booklets. Booklet #1 contained a description of the nature of the experiment, instructions to the experiment, background data, and manipulation checks (see Appendix A). The background data consisted of information about a hypothetical audit client. Although the audit client was a fictional firm, it was modeled after an actual Lynnwood, Washington manufacturing firm.

Booklet #2 of the instrument contained the audit planning cases and response scales (see Appendix B). In these audit planning cases, subjects received information about the client's internal operating environment, the client's internal control structure, and results of analytical procedures. Each of these three factors was varied dichotomously and was operationalized as a repeated-measures variable. Information about the client's external business environment was varied dichotomously as either stable or volatile, and was operationalized as a between-subjects variable. Thus, for any given subject, the information about the client's external business environment did not change from case to case and was presented as part of the unchanging background information in Booklet #1. The manner in which each of the four factors was operationalized produced eight audit planning cases per subject. In order to distinguish between summing- and averaging-type information integration, it was necessary to ask subjects to respond to two additional cases. This raised the total number of audit planning cases per subject from eight to ten.

Booklet #3 contained the internal control evaluation cases, response scales, and the debriefing questionnaire (see Appendix C). In each internal control evaluation case, subjects were presented with information about four separate control procedures. Each control was varied dichotomously as either operating effectively or not. Three of the four controls were operationalized as repeated measures. The fourth control was operationalized as a between-subjects factor. This resulted in eight internal control evaluation cases per subject. As in the audit planning cases, two additional cases were included to allow the testing of summing- versus averaging-type information integration. Each of these two cases contained information about only two, rather than four, internal controls. This resulted in a total of 10 internal control evaluation cases per subject. Each case in both Booklet #2 and Booklet #3 was presented on a separate page. A digram-balanced, latin square design was used to systematically counterbalance the order in which each subject received the stimulus combinations (Wagenaar [1969]). This was done to control for possible practice effects. A cover letter was included which identified the researcher's name, affiliation, and telephone number. Subjects were invited to call the researcher for clarification of any items concerning the experiment which may have been unclear. No subjects requested clarifying information. Subjects were asked to return the booklets to their firm's contact person in an enclosed pre-addressed envelope. The contact person then returned the set of instruments to the researcher.

Auditing students were given the three booklets in their auditing classes near the middle of the ten-week term. The topics of audit risk and internal control had been covered in class lecture. Students, like

their practitioner counterparts, were given instructions concerning the importance of working independently. The booklets were completed in class, with the researcher present.

3.3 COMPONENTS-INTEGRATION STAGE AUDIT JUDGMENTS

One goal of this research is to investigate how auditors integrate the results of tests of individual controls when judging the effectiveness of the internal control structure. In particular, this phase of the research examines the nature of the evidence integration function used by auditors when making this judgment. The first step in identifying the nature of the integration function is to determine whether auditors tend to employ an additive process, or whether their integration process tends to be nonadditive. Previous research on auditors' judgment processes involved in the evaluation of internal controls has found very little evidence of nonadditivity (e.g., Ashton [1974], Ashton and Brown [1980], Ashton and Kramer [1980], and Hamilton and Wright [1982]).

Identifying whether the integration function tends to be additive or nonadditive is an important initial step in determining the form of the functional integration rule used by auditors in the evaluation of the control structure. However, the results of previous research leave some ambiguity in the form of the integration rule. In particular, previous research has not been able to distinguish between summing-versus averaging-type evidence integration. Since the test of additive versus nonadditive integration (traditionally performed through ANOVA's tests of interactive effects) does not distinguish between summing- and averaging-type integration, additional tests are needed to distinguish

between them. The nature of these tests is discussed in section 3.3.1.

To illustrate the difference between summing- and averaging-type information integration, let $R(a)$ represent an individual's judgment based on information set (a) . Also let $R(a + x)$ represent that individual's judgment based on information set (a) and additional information (x) . According to summing-type models of information integration, when $R(a + x) > R(a)$ in terms of the favorableness of the judge's response toward the object being evaluated, then $R(b + x) > R(b)$ for all x . However, according to averaging-type information integration models, if $b > x > a$, then $R(a + x) > R(a)$, but $R(b + x) < R(b)$. What this implies is that adding favorable information to an existing set of information can actually result in a less favorable judgment if the information is integrated in an averaging fashion.

Anderson [1981] points out the practical importance of distinguishing between summing- and averaging-type information integration in his discussion of the manner in which people tend to prepare their vitae when applying for jobs or grants. If an employer integrates information on an applicant's vita in an averaging fashion, it could be disadvantageous for the applicant to add every possible favorable piece of information to his/her vita when preparing it. This is because moderately-favorable information could actually result in a less favorable response by the employer toward the applicant than if the vita had only contained highly-favorable information. Notice that a summing model of information integration would predict the opposite result.

As another example of the difference between summing and averaging integration processes, consider the following. In the evaluation of

internal controls, if a new relevant piece of moderately polarized evidence comes to the auditor's attention and is combined with an existing highly polarized evidence set, the impact of this new piece of evidence on the auditor's evaluation of the internal control structure will depend on whether the auditor integrates evidence by averaging it or by summing it. If the auditor tends to follow a summing integration rule, the addition of moderately favorable (moderately unfavorable) information to a preexisting set of highly favorable (highly unfavorable) evidence will result in an even more favorable (more unfavorable) evaluation of the structure by the auditor. On the other hand, if the auditor integrates evidence via an averaging process, the addition of moderately favorable (moderately unfavorable) information to an existing set of highly favorable (highly unfavorable) evidence will result in a less favorable (less unfavorable) internal control evaluation.

3.3.1 Tests of Summing Versus Averaging Integration of Audit Evidence

Experimentally, a distinction can be made between summing and averaging integration processes. This can be done by adding moderately favorable (unfavorable) information to a highly favorable (unfavorable) information set. If the summing hypothesis is correct, then the response to the information set containing both highly favorable (unfavorable) information and moderately favorable (unfavorable) information should be more favorable (unfavorable) than the response to the highly favorable (highly unfavorable) information set alone. On the other hand, if the averaging hypothesis is correct, then the response to the combined set of favorable information should be less favorable than the response to only the highly favorable information set. Furthermore,

the response to the combined set of unfavorable information should be more favorable than the response to the set containing only highly unfavorable information. One nicety of this test is that it is scale-free and depends on a difference in direction only, not in amount (Anderson [1982]).

The critical test between summing- and averaging-type integration rules is illustrated graphically in Figure 3.1.

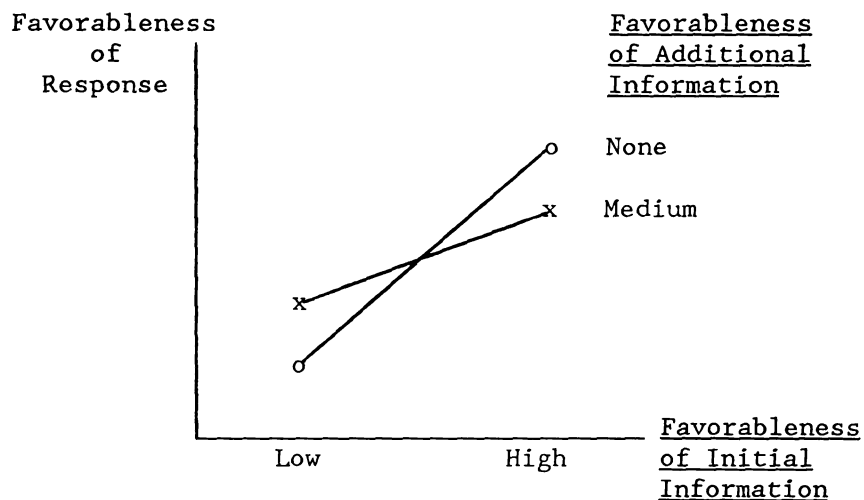


Figure 3.1
Graphical Test of Summing vs. Averaging
Information Integration Processes

Suppose auditors are presented with two sets of highly polarized information -- one set containing very favorable information and the other containing very unfavorable information. The auditors are then asked to make some judgment based on each set. Now suppose moderately polarized information (neither very favorable nor unfavorable) is added to both highly favorable and highly unfavorable information sets and auditors are asked to respond to the resulting two new sets of information (high-medium; low-medium). According to a summing

integration rule, this should change the responses in the same direction, either up or down, depending on whether the medium information tends to be positive or negative. However, if auditors tend to integrate evidence in an averaging fashion, the addition of medium information to highly polarized information should make the responses less extreme than the responses to the highly polarized information alone. If the responses were graphed, an averaging rule would be indicated by a cross-over of the "medium" curve and the "none" curve; a summing rule would produce no such crossover.

3.3.2 Evaluation of Internal Controls -- Independent Variables

To study how auditors integrate the results of tests of individual internal controls into a single evaluation of the effectiveness of the internal control structure, subjects (students, seniors, and managers) were asked to evaluate 10 internal-control-evaluation cases. Subjects' responses to these cases also provided part of the data necessary to test hypothesis eight which deals with the effects of audit experience on audit program planning judgments. Eight of the ten cases were produced from a $2 \times (2 \times 2 \times 2)$ mixed (split-plot) factorial design. Two additional cases were added to permit the testing of summing- versus averaging-type evidence integration. In order to distinguish between averaging- and summing-type evidence integration, not only must the levels of the factors be varied, but the amount of evidence presented to the subjects must also be varied. Accordingly, the two additional cases mentioned above each contained only two factors rather than four as in the other eight cases. Each factor in the design represents an internal control over the sales/cash collections cycle. The levels of each factor were varied dichotomously and indicated whether a particular

control was operating effectively or not. The choice of the sales/cash collections cycle was subjective. However, the audit of this cycle is an important part of most audits and it was believed that most practitioners at the senior level or higher would have experience in this area.

The stimulus combinations were presented to subjects in the form of an internal control questionnaire in which questions had been pre-answered. An answer of "yes" indicates a control is operating as planned, and "no" indicates a weakness in the structure (see Appendix C). The choice of questions was based on a review of the auditing literature and previous research. Each of the questions reflect factors which are considered to be relevant and important in the evaluation of internal controls.

3.3.3 Evaluation of Internal Controls -- Dependent Variable

For each case, subjects were asked to (1) consider the background information about the client presented in Booklet #1, (2) evaluate the information about the client's internal control structure presented in the case, and (3) rate the strength of the internal control structure and indicate their judgment on a 4-inch continuum labeled "extremely weak" on the left end and "adequate to strong" on the right end. This choice of end-point anchors is based on Ashton's [1973] findings that when auditors feel anything is wrong with an internal control structure, its ratings will decrease very sharply from the upper end of the response scale. Ashton found that auditors are much better at differentiating between various degrees of internal control weaknesses than between internal control strengths. Therefore, the response scale did not include labels such as "very strong" and "extremely strong".

One objective of this phase of the experiment is to identify the manner in which auditors integrate tests of individual internal controls when evaluating the effectiveness of the internal control structure. Specifically, an attempt is made to identify whether auditors tend to integrate tests of individual internal control in a summing-type or an averaging-type fashion. However, inasmuch as the author is unaware of any theory on which to form explicit predictions concerning summing-versus averaging-type integration, no specific hypotheses are developed in this regard.

3.4 FACTORS-INTEGRATION STAGE AUDIT JUDGMENTS

In studying the components-integration stage of auditors' judgments discussed in the previous section, the interest is in examining how auditors integrate evidence pertaining to individual internal controls when evaluating the internal control structure. In the study of the factors-integration stage of auditors' judgments, the focus is on how this type of judgment is further utilized and integrated with other audit evidence into a single audit planning judgment concerning the appropriate extent of tests of details of balances to be performed.

3.4.1 Independent Variables

Most of the existing research which has dealt with audit program planning has focused on determining how changes in the effectiveness of individual internal controls affect the planned amount of substantive testing. However, the auditing literature suggests there are several other factors which auditors should consider and evaluate when planning the extent of substantive testing. In this dissertation, four of these

factors have been selected for examination. One of the four factors corresponds to the judgments discussed in the previous section -- the auditor's judgment of the effectiveness of the internal control structure. The other factors selected for examination in this stage are (1) the internal operating environment, other than internal controls, (2) the results of analytical procedures, and (3) the external business environment. These factors were selected for examination subjectively. However, the auditing literature indicates that the auditor's evaluation of each of these factors is important in audit planning (e.g., AICPA [1986] AU 311). Furthermore, each of these factors may potentially affect the auditor's assessment of the components of the audit risk model, and thus affect the extent of tests of details of balances. Other potentially relevant factors were held constant in the background information provided to the subjects.

To test the hypothesized relationships between each of the above factors and auditors' judgments of the appropriate extent of tests of details of balances, subjects were asked to evaluate 10 audit planning cases. These cases were produced from a $2 \times (2 \times 2 \times 2)$ mixed factorial design, with two additional cases included to permit the testing of summing- versus averaging-type evidence integration if warranted. The $2 \times (2 \times 2 \times 2)$ mixed factorial design was used for purposes of power and efficiency. Pure between-subjects designs have the undesirable characteristic of a relatively large error term. This is due to the fact that all individual differences go into that term, making the significance tests of interactions low in statistical power. A pure within-subjects design was not used to keep the number of cases that each subject is asked to evaluate down to a reasonable level. Each

factor in the design corresponds to one of the four factors mentioned above. The between-subjects factor is the external business environment. Internal operating environment, internal control, and analytical procedures were operationalized as within-subjects variables. Each factor was treated as an aggregate variable and the levels of each factor were varied dichotomously as follows: (1) a private, non-publicly held manufacturer of lighting fixtures which has been in business for 66 years and whose profits have been stable with a gradual increasing trend vs. a manufacturer of audio compact disc players which has been in business for 3 years, is expecting to go public within the next year, and whose profits have been irregular with a declining trend; (2) a management that is very cooperative with their auditors, follows conservative accounting practices, has competent employees handling receivables, and whose last year's audit revealed no material misstatements in receivables vs. a management that is uncooperative with their auditors, follows unconservative accounting practices, has rather incompetent employees with a high turnover rate, and whose last year's audit revealed several material misstatements in receivables; (3) an effective internal control structure vs. an internal control structure which has several serious weaknesses; and (4) results of analytical procedures which are consistent with no material misstatements in receivables vs. results of analytical procedures which suggest the presence of material misstatements in receivables (see Appendix B).

3.4.2 Dependent Variables

Joyce [1976] describes an ideal audit program planning measure as one which takes into account (1) the particular audit procedures chosen

to be performed, (2) the extent to which these procedures are performed, and (3) the timing of the audit procedures chosen. The dependent variable used by Joyce (and by Gaumnitz, et. al. [1982]) was able to identify the selection and extent of audit procedures, but did not take into account the timing of the procedures. However, the information on the selection of procedures was found to add no additional useful information in their research. Srinidhi & Vasarhelyi [1986] employed a rating scale which captured the extent and timing of testing, but did not take into account the selection of procedures to perform. In the current study, given the focus on the audit of accounts receivable, and given that several of the substantive audit procedures relevant to testing the accounts receivable balance are normally restricted to year-end or subsequent to year-end (e.g., review of subsequent cash collections), it does not seem beneficial to examine auditors' judgments concerning the timing of testing.

For each case, auditors were asked to indicate their judgments regarding the total number of hours they would plan for tests of details of the accounts receivable balance. In addition, subjects were asked to provide a second response to each case. This second response was elicited by asking the subjects to indicate graphically their planned extent of testing of receivables by placing a slash mark on a four-inch line, anchored with "no testing of receivables" and "extensive testing of receivables". The purpose of this second response measure is to provide a dependent measure with less between-subject response variability than is likely to be present in the "number of hours" dependent variable. Excessive between-subject response variability in the reported "number of hours" could mask the presence of significant

main effects or interactions of factors when analyzing the subjects' responses. In addition, a participant in a pilot test of the instrument pointed out that a decision to change from confirming 10% of the accounts receivable to confirming 100% of the receivables may not necessarily require much more time if the confirmations are prepared via a computer.

A summary of the dependent and independent variables involved with each set of cases is presented in Figure 3.2.

	<u>Independent Variables</u>	<u>Dependent Variable</u>
<u>Internal Control Cases</u>	Effectiveness of individual internal controls is varied dichotomously.	Scale rating of internal control effectiveness.
<u>Audit Planning Cases</u>	Internal Operating Environment, Internal Control Structure, Results of Analytical Procedures, and External Business Environment are each varied dichotomously.	(1) Budgeted audit hours for accounts receivable, (2) Scale rating of the extent of testing of details of the accounts rec'ble balance.

Figure 3.2
Summary of Variables

3.4.3. Tests of Hypotheses

Seven hypotheses were developed in chapter two to represent expected relationships between the independent variables above and auditors' judgments of the appropriate extent of tests of details of balances. These hypotheses were tested in a complete factorial design, which allowed investigation of the effects of multiple factors simultaneously. Not only can the *separate* effects of each of the factors be studied, but the effects of *combinations* of factors can also

be examined. An alternative approach would be to conduct a series of single-factor experiments in which only one variable is manipulated at a time, with the other factors held constant. However, the factorial design is more appropriate for addressing the issues of this research. Three hypotheses concern the combined effects of two factors. A series of single-factor experiments would not provide enough information to study these combined effects. As stated by Cochran and Cox ([1957], pg 151):

... the single-factor approach is likely to provide only a number of disconnected pieces of information that cannot easily be put together. In order to conduct an experiment on a single factor A, some decision must be made about the levels of other factors B, C, D, say, that are to be used in the experiment... . The experiment reveals the effects of A for this particular combination of B, C, and D, but no information is provided for predicting the effects of A with any other combination of B, C, and D. With a factorial approach, on the other hand, the effects of A are examined for every combination of B, C, and D that is included in the experiment. Thus a great deal of information is accumulated both about the effects of the factors and about their inter-relationships.

Several previous studies of auditors' judgments employ *fractional* factorial designs. Fractional factorial designs reduce the number of cases per subject as compared to a complete factorial design. For example, a complete factorial experiment in which four variables are manipulated dichotomously requires 16 cases. A $1/2$ fractional factorial design requires only 8 cases. However, this reduction in the size of the experiment brings with it certain disadvantages. One disadvantage is the inability to estimate higher-order interactions. In a $1/2$ replicate of a 2^4 design, the four-factor interaction cannot be estimated (Winer [1962], pg 449). Another disadvantage of fractional factorial designs is that the effects due to higher-order interactions

are confounded with lower-order interactions and main effects. In a $1/2$ replicate of a 2^4 factorial design, the main effects are confounded with three-way interactions, and two-way interactions are confounded with other two-way interactions (Cox [1958], pg. 253). For instance, the effect due to factor A is confounded, or "aliased", with the BCD interaction, and the effect due to the AB interaction is confounded ("aliased") with the CD interaction.

This confounding is not too serious a problem in higher-order designs if one can assume the higher-order interactions are negligible and if an appropriate design is chosen so that main effects are not aliased with lower-order interactions. However, in this study, higher-order interactions cannot be assumed negligible. Furthermore, it would not be desirable to allow two-way interactions to be confounded with other two-way interactions. Accordingly, this study does not use a fractional factorial design. Rather, a complete factorial design is employed in which each combination of levels of every factor is used.

The $2 \times (2^3)$ design used in this research permits the estimation of all main effects and all interactions. A significant main effect (or set of significant simple main effects) for a given factor indicates that the auditors' responses vary systematically with changes in the levels of that factor. In other words, that factor makes a difference in the auditors' judgments. This research hypothesized that each of the four factors in the experimental design plays a significant role in audit planning judgments (hypotheses one through four). Equivalently, significant main effects for each factor were hypothesized. For factors which interact with other factors, significant simple main effects across all levels of the other factors with which they interact were

hypothesized.

Hypotheses five, six and seven predicted specific two-way interactions between certain factors. Accordingly, the evaluation of subjects' responses focused initially on interaction effects. As explained by Keppel [1982]:

A test for interaction usually represents a logical first step in the analysis of a[n] ... experiment in the sense that the outcome of this test generally will influence the nature of the analyses that follow. ... the absence of statistically significant interaction usually means that any subsequent analyses will generally focus on the marginal means [main effects] rather than on the individual treatment means [simple effects]. Stated another way, we can describe and analyze the effects of one of the independent variables without considering the specific levels of the other independent variable. ... [however,] any analyses conducted after the establishment of a significant interaction will tend to concentrate on the individual treatment means [simple effects] rather than on the overall marginal means [main effects] (pp. 179, 209).

A significant two-way interaction between two factors suggests that the effect of one factor on auditors' judgments depends on the level of the other factor. For example, in the matrix below, the effect of factor A depends on the level of factor B:

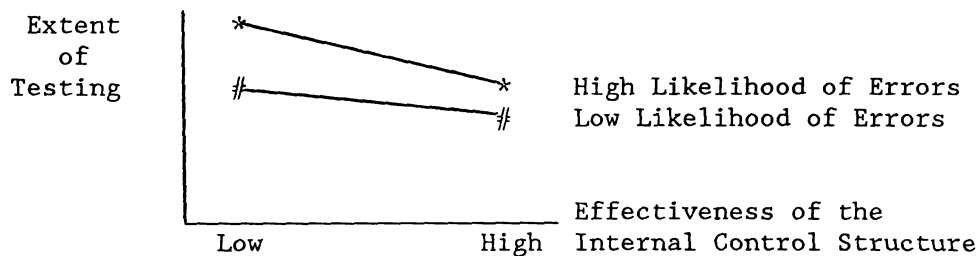
		Factor A	
		High	Low
Factor B	High	9	9
	Low	11	5

When factor B is high, factor A has no effect on the response. But when factor B is low, a change in factor A from high to low reduces the response from 11 to 5. Alternatively, the effect of factor B depends on the particular level of factor A.

Experts often claim they use information in a configural (interactive) fashion. However, empirical research has seldom

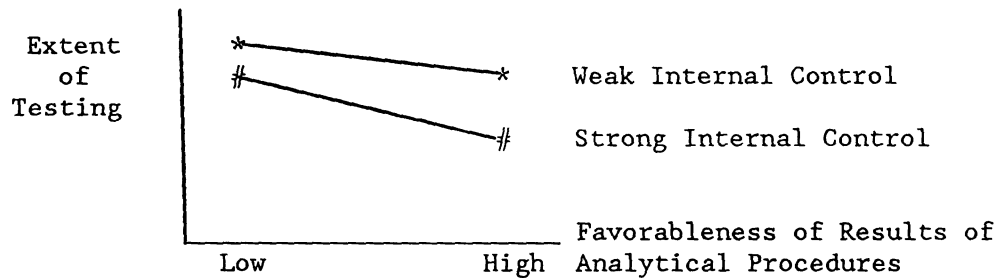
substantiated this claim. The auditing literature suggests that auditors should use certain information in a configural fashion as well. A goal of this study is to address this issue in the area of audit planning.

The fifth hypothesis suggests that an increase in the effectiveness of the internal control structure will have a larger effect on auditors' judgments when the client's internal operating environment is such that the likelihood of material misstatements occurring is high than when it is low. Thus, an ordinal interaction of the following form is expected between internal control and the internal operating environment:

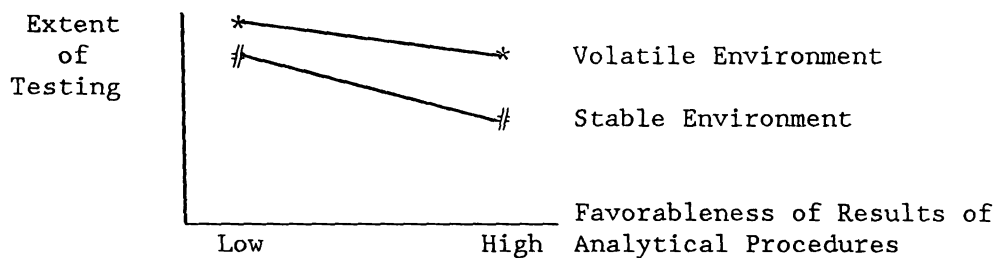


Whether the subjects in this experiment tend to use the factors in the hypothesized configural manner is examined by testing the F-ratios for the hypothesized two-factor interactions. The calculation of the F-ratios is based on ANOVA.

Hypotheses six and seven are tested in a similar manner. The sixth hypothesis predicts an ordinal interaction between the results of analytical procedures and internal control effectiveness of the following form:



The seventh hypothesis predicts the following type of ordinal interaction between the results of analytical procedures and the external business environment:



Hypotheses six and seven are tested by using ANOVA to determine the significance of the F-ratios for the Analytical Procedures X Internal Control interaction and Analytical Procedures X Business Environment interaction, respectively.

In addition to calculating the F-ratios for each factor and two-way interactions, the relative strength of the effects of each factor and interaction on the subjects' judgments is assessed by calculating the eta-squared index for each factor and interaction. Eta-squared, in this context, represents the proportion of the variance in auditors' program planning judgments that is accounted for by each factor and interaction of factors.

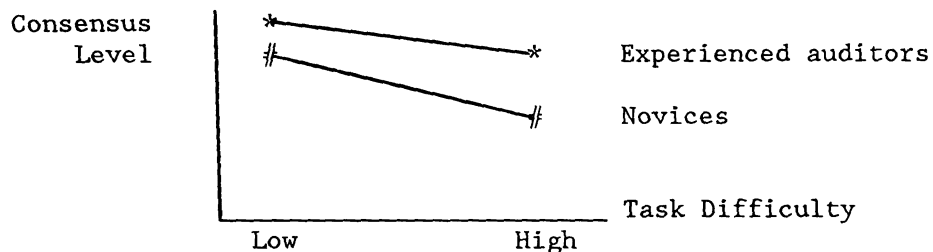
3.5 EFFECTS OF EXPERIENCE AND AUDIT TASK

To examine how audit experience affects auditors' performance in tasks which require professional judgment, three groups of subjects

differing in their levels of experience were asked to perform two audit tasks differing in difficulty. The two tasks chosen for examination are the same tasks discussed in sections 3.2 and 3.3 -- internal control evaluation and audit planning. The effects of the auditors' levels of experience on their ability to correctly perform these two audit tasks are examined as follows.

Accounting judgment studies which have utilized the Brunswik lens model framework have typically focused on only the right side of the lens, as there is frequently no knowledge of the "true state" on the left side of the lens with which to assess judgmental accuracy. Accordingly, such studies have tended to use consensus as a measure of judgmental performance. The current study will also use consensus as a measure of performance.

This phase of the research focuses on levels of judgmental agreement (consensus) between auditors. Hypothesis eight predicts that there are larger differences between the consensus levels (a surrogate for accuracy) of audit practitioners and auditing students when performing a task which requires higher levels of expertise than when performing a task requiring little expertise. Two methods were chosen to measure judgmental agreement -- the Pearson product-moment correlation coefficient, and mean absolute differences. The mean pairwise inter-subject agreement is calculated separately for the set of practitioners and for the set of students and constitutes the dependent variable in the analysis; experience level and level of audit procedure difficulty are the independent variables. Hypothesis eight predicts an ordinal interaction between experience level and audit task difficulty of the following form:



ANOVA is used to test the significance of the interaction.

In addition to examining the effects of experience on consensus, three possible sources of inter-auditor judgmental disagreement are investigated. The social judgment version of the lens model identifies three particular factors which may influence the extent to which the judgments of two individuals agree. These three factors are cue weighting agreement, judgment consistency, and extent of configurality (Hammond, Stewart, Brehmer, and Steinman [1975]; Bonner [1988]). If two individuals agree on how individual cues should be weighted when forming a particular judgment, then the predictions of the optimal (based on the least-squares criterion) regression models of individuals 1 and 2 should be similar. For reasons described in Chapter Four, the extent of similarity between the models is measured, in this study, by mean absolute differences. If an individual is consistent in the manner in which he/she forms a particular type of judgment, then that individual's judgments should consistently be similar to judgments predicted by the optimal regression model of that individual. Again, mean absolute differences is chosen as the measure of similarity in these judgments. Finally, if two individuals have similar degrees of configurality in their judgment processes, then regression models of those two individuals' judgments should have similar residual unexplained variances. Mean absolute differences between residual variances is used

as the measure of the extent of agreement on configural cue processing in this study.

The above discussion suggests that the degree of consensus between the judgments of two individuals depends on (1) the extent to which they agree on how cues should be weighted (G_{MAD}), (2) the consistency with which each of the individuals makes his/her judgments (R_{MAD}), and (3) their agreement on configural cue utilization (C_{MAD}). The objective at this stage is to assess whether inter-auditor judgmental disagreement is due mainly to differences in cue-weighting agreement, differences in judgment consistency, or differences in the degree of configurality in the judgment process. To address this issue, a series of 2 X 2 ANOVAs are run with Experience Level and Task Difficulty as the independent variables and G_{MAD} , R_{MAD} , and C_{MAD} as the dependent variables. An examination of the Experience Level X Task Difficulty interaction in each of the ANOVAs is used to identify which of the three components of consensus (G_{MAD} , R_{MAD} , or C_{MAD}) is most responsible for inter-auditor judgmental differences.

The analyses above focus on comparisons of the judgments of audit practitioners and auditing students. In order to gain further insight into how and where auditors develop the knowledge and expertise required for professional audit program planning judgments, the "practitioners" group of subjects is decomposed into two groups -- seniors and managers. Analyses similar to those discussed above are performed on comparisons of the judgments of students with those of seniors, and judgments of seniors with those of managers. If auditors develop the relevant expertise within the first three or four years after graduation from college, an *insignificant* Experience-level X Procedure Difficulty-level

interaction would be expected when the dependent variable is judgmental agreement (consensus) of all pairs of *seniors* and all pairs of *managers*, and a *significant* Experience-level X Procedure Difficulty-level interaction based on the judgments of *students* and *seniors* would be expected. On the other hand, if audit program planning expertise is developed later, *significance* in the experience X difficulty interaction based on the judgments of *seniors* and *managers* would be expected, and *insignificance* in the experience X difficulty interaction based on the judgments of *students* and *seniors* would be expected.

3.6 PILOT TESTS

Three pilot tests of the experiment were performed. The initial pilot test involved 42 accounting students enrolled in their first undergraduate auditing class. The results of this pilot test resulted in minor changes to the instructions in Booklet #1 and to the wording of one of the manipulation checks. The second pilot test involved an auditing professor at Michigan State University. Three changes resulted from this pilot test. One change involved the addition of clarifying instructions to the beginning of Booklet #2 and Booklet #3. A second change involved moving the manipulation checks from Booklet #2 to the end of Booklet #1. This was done to "clean up" the presentation of the experimental cases in Booklet #2. The third change involved moving the two "short" cases in both Booklets #2 and #3 from the end of the booklets to the beginning. This was done to reduce the risk of "carry-over" effects in which information presented in the long cases might influence the judgments made in the short cases. In addition, the participant in the second pilot test suggested performing a third pilot

test involving a practitioner in the CPA firm that was to participate in the experiment. Since only one CPA firm was to participate in the experiment, the intent of this third pilot test was to alter the instrument so as to be consistent with the methodology and terminology used by that firm. Several changes in the wording of the instructions, background information, manipulation checks, and debriefing questionnaire emerged from this pilot test. These changes were made to be consistent with the terminology used by the firm. These changes were deemed desirable so as to minimize the risk of the practitioner subjects misunderstanding any part of the instrument. No changes were made to the actual cases.

3.7 SUBJECTS

Twenty seniors and 19 managers provided responses to the audit planning experiment. The same 20 seniors and 18 of the 19 managers also provided responses to the internal control evaluation experiment. This reflects a response rate to both experiments of 100% for seniors, while managers had a response rate of 95% to the audit planning experiment, and 90% to the internal control evaluation experiment. All seniors and 15 of the managers were from the Detroit office of their firm. The remaining 4 managers worked in the Grand Rapids, Michigan office of their firm. All practitioner subjects were CPAs. The mean auditing experience of managers was 7.3 years, and ranged from 5 to 11 years. For seniors, the mean auditing experience was 3.4 years, and ranged from 3 to 4 years.

Thirty-seven undergraduate accounting students who were enrolled in their first undergraduate auditing course completed both experiments.

Responses from four students with auditing experience from internships were discarded. Thirteen of the remaining students' responses were systematically eliminated in order to arrive at a final student sample size of twenty. These 13 students' responses were eliminated in a manner which maintained the digram-balanced nature of the experimental design, and which maintained an equal number of student subjects in each external-business-environment condition. A summary of the number of subjects at each level of audit experience whose responses were used in the data analysis is presented in Table 3.1. Results of the experiments are presented in the next chapter.

Table 3.1
Number of Subjects at each Experience Level

	Audit Planning Cases	Internal Control Evaluation Cases
Students	20	20
Seniors	20	20
Managers	19	18
Totals	59	58

CHAPTER 4 - RESULTS

4.1 INTRODUCTION

This research addresses the issues of (1) the manner in which auditors use and integrate information in audit planning, (2) the effects of professional experience on auditors' audit planning and internal control evaluation judgments, and (3) the nature of the evidence integration processes employed by auditors when evaluating the internal control structure. In Chapter 2, seven hypotheses were developed regarding the manner in which (1) the client's internal operating environment, (2) the client's internal control structure, (3) the results of analytical procedures, and (4) the client's external business environment affect auditors' judgments of the appropriate extent of tests of details of balance. In addition, a hypothesis was developed regarding the relationship between professional experience and judgmental agreement between auditors. An experiment was performed to examine these issues. Practicing auditors and accounting students were asked to respond to a series of audit planning and internal control evaluation cases.

In the current chapter, the results of the experiment are analyzed. Each of the hypotheses developed in Chapter 2 is tested and evaluated. The relative weighting, or importance, of the internal operating environment, the control structure, the results of analytical procedures, and the client's business environment on auditors' extent-of-tests-of-details judgments is investigated also.

As part of the evaluation of hypothesis eight regarding the effects of professional experience on judgmental agreement between

auditors, potential sources of judgmental disagreement are examined. In addition, an investigation is made of how and where professional expertise relating to audit program planning is developed.

Finally, the manner in which auditors integrate tests of individual internal controls when evaluating the effectiveness of the internal control structure is examined.

4.2 USE AND INTEGRATION OF EVIDENCE IN AUDIT PLANNING

4.2.1 Manipulation Checks

Manipulation checks were performed to determine whether the factors in the experiment were operationalized as planned. The manipulation checks were included near the end of booklet #1 (Appendix A). Manipulation checks for internal operating environment, internal control, and analytical procedures were all statistically significant ($P < .001$). External business environment was marginally significant ($F = 3.43$, $P = .069$). The manipulation check for external environment was performed by asking subjects to indicate their acceptable level of overall risk for the current year's audit of the company described in the background information. Subjects responded by placing a slash-mark on a nine-point scale anchored "Low (e.g., 1%)" and "High (e.g., 9%)". Given the marginal significance of the external-environment manipulation check, separate checks were performed for practitioners and students. For practitioners, external operating environment was significant ($F = 7.05$, $P = .011$). For students, however, external environment was not significant ($F < 1.0$). Although students apparently had a difficult time judging how changes in the environment affect the acceptable level of overall risk, their planned extent-of-testing ratings, the variable

of interest in this study, was found to vary systematically with changes in the external environment. This indicates that students did indeed perceive differences across the external-environment treatments.

Therefore, all factors appear to have been operationalized as intended.

4.2.2 Budgeted Audit Hours

Hypothesis one predicts a *direct* relationship between auditors' planned extent of tests of details of balances and conditions in the clients' internal operating environment which increase the likelihood of the occurrence of material misstatements. Hypothesis two predicts an *inverse* relationship between the planned extent of tests of details and the effectiveness of the internal control structure. Hypothesis three predicts an *inverse* relationship between the planned extent of tests of details and the favorableness of the results of analytical procedures. Hypothesis four predicts a *direct* relationship between the planned extent of tests of details and the degree of volatility in the client's external business environment. Two methods were used in the experiment to measure subjects' planned extent of tests of details judgments. One method involved asking subjects to indicate how many hours they would budget for tests of details of the accounts receivable balance. The second method involved asking subjects to indicate graphically their planned extent of tests of details of the accounts receivable balance by placing a slash mark on a four-inch line anchored "no testing of receivables" on the left and "extensive testing of receivables" on the right. Subjects' budgeted audit hours responses are analyzed in this section of the chapter. Subjects' extent-of-testing ratings are analyzed in Section 4.2.3.

Cell means and standard deviations for budgeted audit hours (averaged across experience levels), and cell means across experience levels are presented in Table 4.1. In all cases, average budgeted hours were greater when the internal operating environment was unfavorable than when it was favorable, *ceteris paribus*. Similarly, average budgeted hours were greater when internal control was weak than when it was strong, and when the results of analytical procedures were favorable than when they were unfavorable. This pattern of results is consistent with hypotheses one, two, and three. However, contrary to hypothesis four, budgeted audit hours were not consistently greater in the volatile external business environment than in the stable external business environment. For instance, the bottom left-most cell in Table 4.1 shows average budgeted hours of 113.11 in the unfavorable internal operating environment / weak internal control / unfavorable analytical procedures / volatile external environment case. However, the equivalent cell in the stable external environment case shows average budgeted audit hours of 135.44. One possible explanation for this occurrence may be that in a stable business environment, the auditor has more confidence in the results of his/her analytical procedures (see SAS 56) than in a more volatile environment. Accordingly, in a stable environment, the auditor places more weight on negative analytical results than when the environment is volatile, resulting in the observed increase in budgeted hours as one moves from a volatile environment to a stable environment. This same pattern of results is found in column three, rows one and two, of Table 4.1 where budgeted hours for the favorable internal operating environment / strong internal control structure / unfavorable analytical procedures / volatile external environment case is 64.44, and the

Table 4.1
Budgeted Audit Hours
Cell Means and Standard Deviations

Panel A
Collapsed Across Experience Level

		Internal Operating Environment				
		<u>Unfavorable</u>		<u>Favorable</u>		
		Results of Analytical Procedures		Results of Analytical Procedures		
		<u>Unfavor- able</u>	<u>Favor- able</u>	<u>Unfavor- able</u>	<u>Favor- able</u>	<u>Averages</u>
Strong Internal Control Structure	Stable	81.67	67.33	64.63	32.56	61.55
	External	(150.40)	(148.11)	(114.72)	(36.14)	
	Environment					65.16
	Volatile	95.74	69.48	64.44	45.37	68.76
	External	(181.32)	(111.74)	(121.55)	(80.83)	
	Environment					
Weak Internal Control Structure	Stable	135.44	75.30	78.74	68.00	89.37
	External	(376.34)	(148.56)	(151.47)	(148.68)	
	Environment					90.04
	Volatile	113.11	96.93	85.33	67.41	90.70
	External	(213.21)	(181.36)	(150.83)	(109.74)	
	Environment					
<u>Averages</u>		106.49	77.26	73.29	53.33	
		91.88		63.31		77.59

Table 4.1 (cont'd.)

Panel B
Grouped by Experience Level

<u>Students</u>	<u>Seniors</u>	<u>Managers</u>
166.41 (265.06)	26.77 (13.26)	39.60 (26.05)

equivalent cell for the stable external environment case shows budgeted hours of 64.63.

The average cell means and standard deviations for students', seniors', and managers' responses are also presented in Table 4.1. As a group, students felt it was necessary to spend much more time on the audit of receivables (average budgeted hours of 166.41) than did either the seniors (average budgeted hours of 26.77) or the managers (average budgeted hours of 39.60). In addition, students seemed to disagree with one another regarding the appropriate number of audit hours to a greater extent than did seniors or managers. This is evident in the standard deviations of the responses of the three groups of subjects shown in parentheses in Table 4.1, as well as in the range of responses. For students, budgeted hours ranged from 8 to 2000 hours, while seniors' and managers' responses ranged from 6 to 80 hours and 10 to 140 hours, respectively. Apparently, students are very uncertain as to how many hours are necessary to complete an audit of a particular account such as receivables. This result seems reasonable since none of the students had ever been involved in an actual audit and the time required to perform various phases of an audit is not addressed in the classroom.

To test the hypothesized relationships between budgeted audit hours and each of the four factors, a $3 \times 2 \times (2^3)$ ANOVA was performed in which subjects' budgeted audit hours was the dependent variable, and subjects' experience level (factor A), the client's external business environment (factor B), the client's internal operating environment (factor C), the client's internal control structure (factor D), and the results of the auditor's analytical procedures (factor E) were the independent variables. The ANOVA results are presented in Table 4.2.¹

Table 4.2
Budgeted Audit Hours
ANOVA Table

Source	SS	df	MS	F	P	Eta ²
Experience Level (A)	1715757.39	2	857878.697	5.32	.008 **	.144
Envirnmmntl Volatility (B)	1967.79	1	1967.787	0.01		.000
A x B	4372.56	2	2186.280	0.01		.000
S/AB	7745715.06	48	161369.064			
Intnl. Oper. Environ. (C)	88122.45	1	88122.454	7.51	.008 **	.007
A x C	78659.67	2	39329.836	3.35	.043 *	.007
B x C	16.33	1	16.333	0.00		.000
A x B x C	850.60	2	425.299	0.04		.000
C x S/AB	563222.94	48	11733.811			
Internal Control (D)	66851.56	1	66851.565	6.72	.012 *	.006
A x D	74978.67	2	37489.336	3.77	.030 *	.006
B x D	936.33	1	936.333	0.09		.000
A x B x D	4682.76	2	2341.382	0.24		.000
D x S/AB	477252.67	48	9942.764			
C x D	334.26	1	334.259	0.43		.000
A x C x D	1599.17	2	799.586	1.04		.000
B x C x D	178.90	1	178.898	0.23		.000
A x B x C x D	283.95	2	141.975	0.18		.000
C x D x S/AB	37057.22	48	772.025			
Analytical Procedures (E)	65317.93	1	65317.926	8.89	.004 **	.006
A x E	86789.45	2	43394.725	5.90	.005 **	.007
B x E	2417.79	1	2417.787	0.33		.000
A x B x E	5675.28	2	2837.641	0.39		.001
E x S/AB	352786.56	48	7349.720			
C x E	2324.08	1	2324.083	1.46	.232	.000
A x C x E	3106.10	2	1553.049	0.98		.000
B x C x E	1160.33	1	1160.333	0.73		.000
A x B x C x E	1574.76	2	787.382	0.49		.000
C x E x S/AB	76416.22	48	1592.005			
D x E	296.68	1	296.676	0.16		.000
A x D x E	1261.17	2	630.586	0.34		.000
B x D x E	2151.15	1	2151.148	1.15	.288	.000
A x B x D x E	4216.12	2	2108.058	1.13	.331	.000
D x E x S/AB	89481.39	48	1864.196			
C x D x E	5720.33	1	5720.333	0.96		.001
A x C x D x E	7989.04	2	3994.521	0.67		.001
B x C x D x E	9766.01	1	9766.009	1.64	.206	.001
A x B x C x D x E	17776.12	2	8888.058	1.49	.234	.002
C x D x E x S/AB	285755.50	48	5953.240			
Totals	11884822.29	431				

* .01<P<.05

** P<.01

In the absence of significant interactions involving factors C, D, and E, hypotheses one through three would be supported by the observed significant main effects for those factors. However, significant interactions were found between experience level and the internal operating environment, experience level and internal control, and experience level and the results of analytical procedures. These interactions, along with individual treatment means (simple effects), need to be further analyzed before conclusions can be drawn regarding hypotheses one, two, or three. The nature of these interactions is presented in Figures 4.1 through 4.3. As shown in Figure 4.1, while the magnitude of the effect of internal operating environment on budgeted hours differed across experience level, the nature of the effect remained consistent across experience level with more audit hours being budgeted in the unfavorable internal environment than in the favorable internal environment. This produced an ordinal interaction between audit experience and internal operating environment. Similar patterns of results are observed for internal control (Figure 4.2) and analytical procedures (Figure 4.3).

The analyses of interactions and simple effects proceeded by creating a set of smaller factorial designs. This approach to analyzing interactions is referred to by Keppel as the analysis of "interaction contrasts" (Keppel [1982], pg. 227). Three sets of 2×2 ANOVAs were formed for each of the three significant 3×2 interactions, the results of which are presented in Tables 4.3, 4.4, and 4.5. The first set involved comparisons of students' and seniors' responses at both levels of the second variable involved in the interaction. Sets two and three were similar, but compared seniors' and managers' responses, and

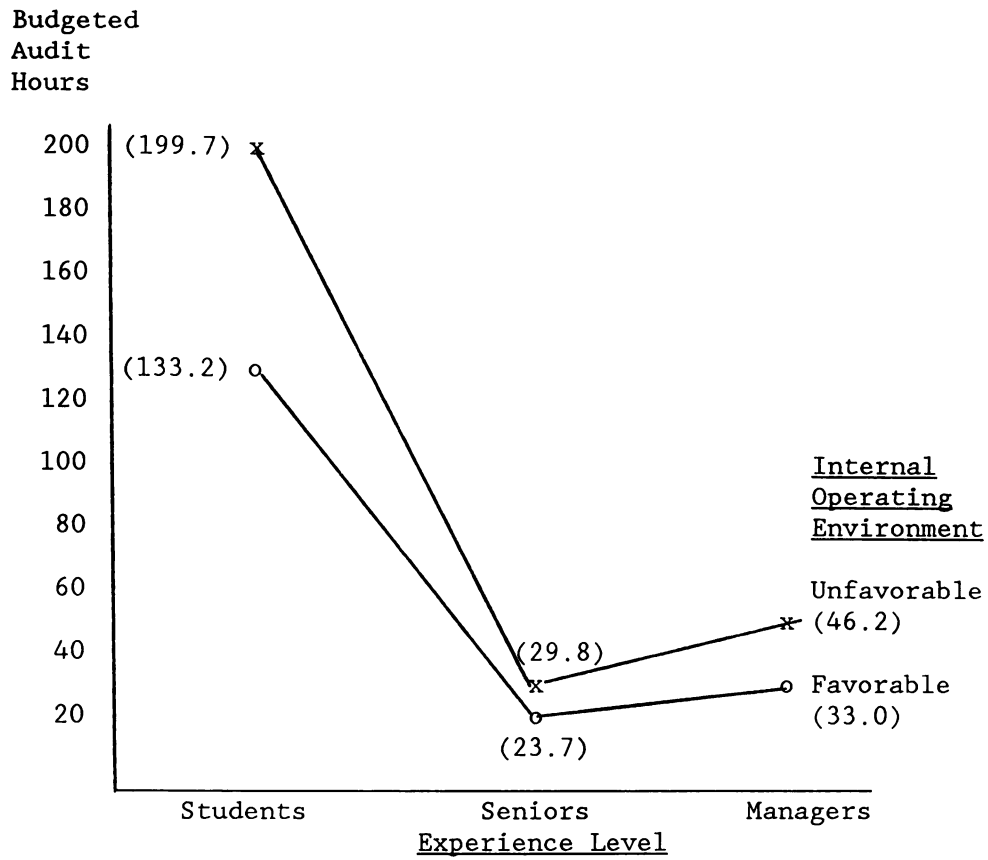


Figure 4.1
Budgeted Audit Hours
Interaction Between Experience Level
and Internal Operating Environment (A x C)

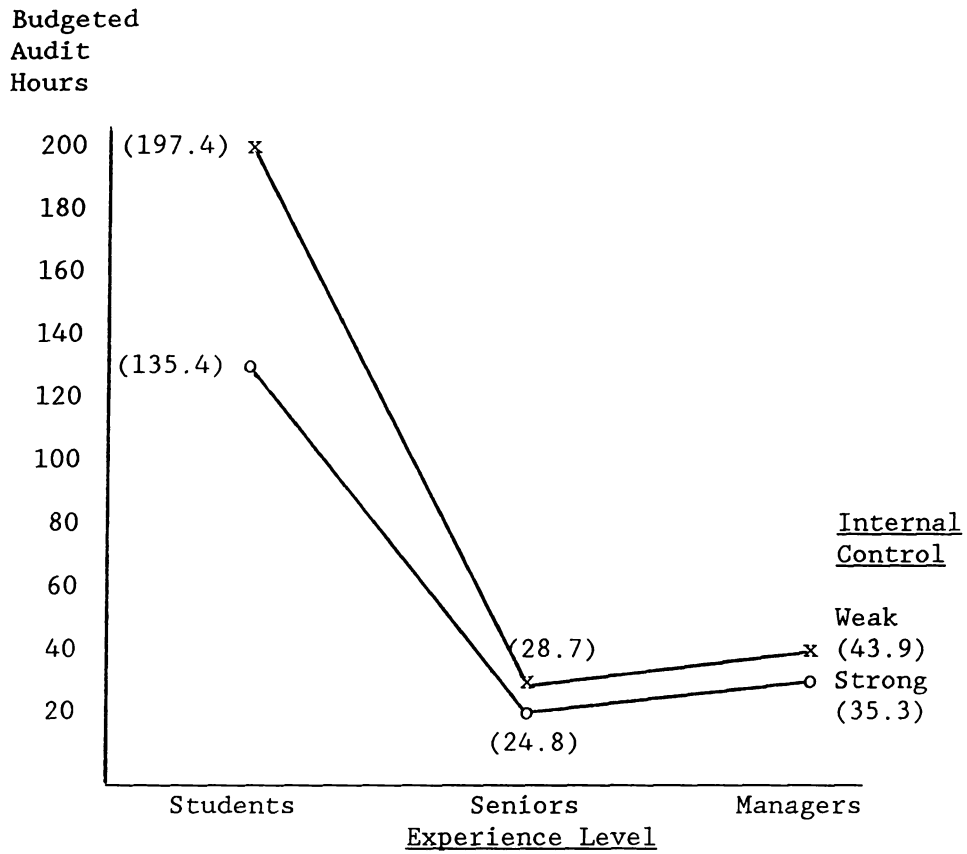


Figure 4.2
Budgeted Audit Hours
Interaction Between Experience Level
and Internal Control (A x D)

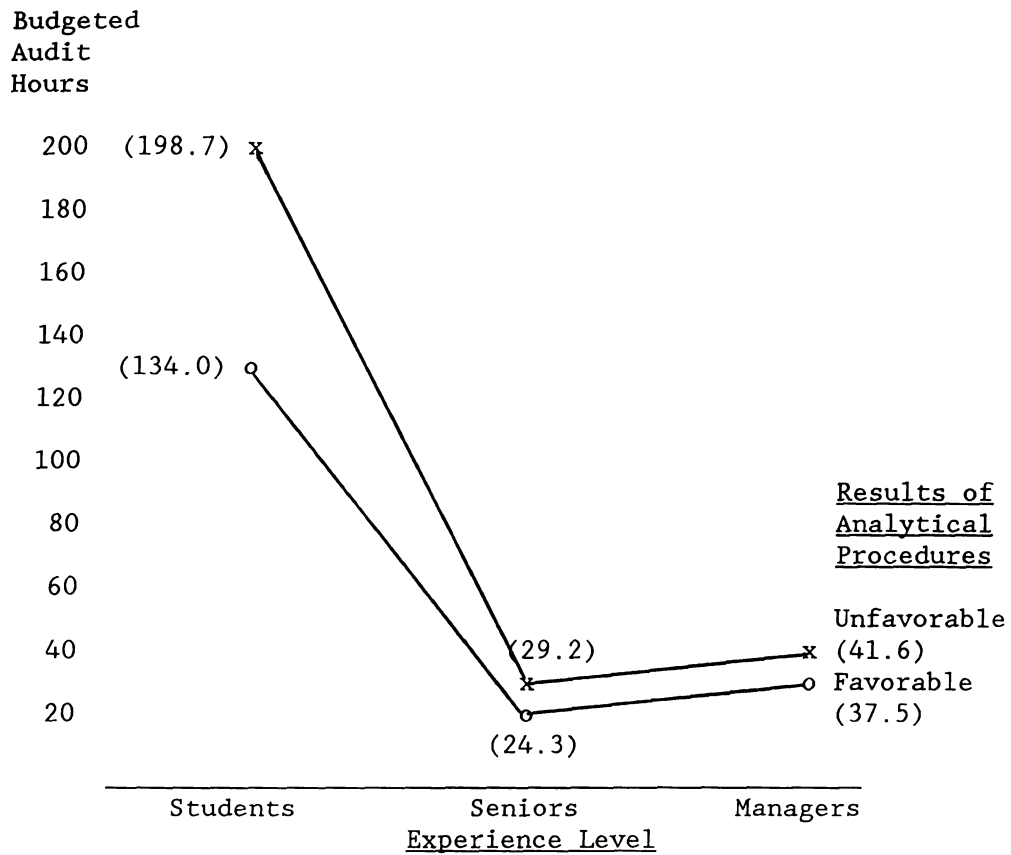


Figure 4.3
Budgeted Audit Hours
Interaction Between Experience Level
and Results of Analytical Procedures (A x E)

students' and managers' responses, respectively. As indicated in Tables 4.3 and 4.4, each of the A x C interactions and A x D interactions based on the smaller 2 x 2 factorials was statistically significant². This suggests that the effects exerted by the internal operating environment (factor C) and internal controls (factor D) on students' judgments of the appropriate number of audit hours to be budgeted are different than the effects on seniors' or managers' judgments. Similarly, the effects on the judgments of seniors differ from those on the judgments of managers. An examination of Figure 4.1 reveals that the judgments of students are influenced to a greater extent by changes in the internal operating environment than are judgments of seniors or managers, and managers are influenced by those same changes to a greater extent than seniors. Figures 4.2 and 4.3 reveal the same pattern of effects for changes in internal control and results of analytical procedures, respectively. However, as indicated by the nonsignificant experience-level x analytical-procedures interaction in Panel B of Table 4.5, the effects of analytical procedures on judgments of seniors and managers are roughly the same.

To complete the testing of hypotheses one, two, and three based on budgeted audit hours responses, the A x C, A x D and A x E interactions were further decomposed by examining the simple effects of internal operating environment, internal controls, and analytical procedures at each of the three levels of experience. All nine of these simple effects were statistically significant at an alpha level of .001 or lower. Given the significance of each of the simple effects of factors C, D, and E across the three levels of experience, and given the ordinal nature of the A x C, A x D, and A x E interactions, hypotheses one, two,

Table 4.3
 Budgeted Audit Hours
 Experience Level x Internal Operating Environment (A x C)
 Interaction Contrasts

Panel A: Students & Seniors

Source	SS	df	MS	F	P
Experience Level (A) (Students & Seniors)	1403929.39	1	1403929.389	22.39	.000 **
S/A	8902273.26	142	62692.065		
Intrnl Oper. Environ. (C)	94830.13	1	94830.125	13.34	.000 **
A x C	65763.56	1	65763.556	9.25	.002 **
C x S/A	1009163.32	142	7106.784		

Panel B: Seniors & Managers

Source	SS	df	MS	F	P
Experience Level (A) (Seniors & Managers)	11845.17	1	11845.170	15.67	.000 **
S/A	107339.58	142	755.913		
Intrnl Oper. Environ. (C)	6622.09	1	6622.087	128.25	.000 **
A x C	892.53	1	892.531	17.29	.000 **
C x S/A	7331.88	142	51.633		

Panel C: Students & Managers

Source	SS	df	MS	F	P
Experience Level (A) (Students & Managers)	1157861.53	1	1157861.531	18.34	.000 **
S/A	8964688.97	142	63131.612		
Intrnl Oper. Environ. (C)	114122.53	1	114122.531	15.99	.000 **
A x C	51333.42	1	51333.420	7.19	.008 **
C x S/A	1013768.55	142	7139.215		

** P<.01

Table 4.4
Budgeted Audit Hours
Experience Level x Internal Control (A x D)
Interaction Contrasts

Panel A: Students & Seniors

Source	SS	df	MS	F	P
Experience Level (A) (Students & Seniors)	1403929.39	1	1403929.389	22.17	.000 **
S/A	8991823.26	142	63322.699		
Internal Control (D)	78342.01	1	78342.014	11.82	.000 **
A x D	60784.22	1	60784.222	9.17	.002 **
D x S/A	941080.76	142	6627.329		

Panel B: Seniors & Managers

Source	SS	df	MS	F	P
Experience Level (A) (Seniors & Managers)	11845.17	1	11845.170	14.93	.000 **
S/A	112690.58	142	793.596		
Internal Control (D)	2856.42	1	2856.420	65.05	.000 **
A x D	403.75	1	403.753	9.19	.002 **
D x S/A	6235.33	142	43.911		

Panel C: Students & Managers

Source	SS	df	MS	F	P
Experience Level (A) (Students & Managers)	1157861.53	1	1157861.531	18.15	.000 **
S/A	9056477.97	142	63778.014		
Internal Control (D)	89994.03	1	89994.031	13.51	.000 **
A x D	51280.03	1	51280.031	7.70	.006 **
D x S/A	946161.44	142	6663.109		

** P<.01

Table 4.5
Budgeted Audit Hours
Experience Level x Results of Analytical Procedures (A x E)
Interaction Contrasts

Panel A: Students & Seniors

Source	SS	df	MS	F	P
Experience Level (A) (Students & Seniors)	1403929.39	1	1403929.389	22.02	.000 **
S/A	9052490.26	142	63749.931		
Analyt'l Procedures (E)	87153.13	1	87153.125	14.26	.000 **
A x E	64320.89	1	64320.889	10.52	.001 **
E x S/A	868065.99	142	6113.141		

Panel B: Seniors & Managers

Source	SS	df	MS	F	P
Experience Level (A) (Seniors & Managers)	11845.17	1	11845.170	14.33	.000 **
S/A	117359.58	142	826.476		
Analyt'l Procedures (E)	1489.67	1	1489.670	63.57	.000 **
A x E	9.03	1	9.031	0.39	N/S
E x S/A	3327.80	142	23.435		

Panel C: Students & Managers

Source	SS	df	MS	F	P
Experience Level (A) (Students & Managers)	1157861.53	1	1157861.531	18.02	.000 **
S/A	9124315.97	142	64255.746		
Analyt'l Procedures (E)	85387.78	1	85387.781	13.96	.000 **
A x E	65854.25	1	65854.253	10.77	.001 **
E x S/A	868355.47	142	6115.179		

** P<.01

and three are supported when auditors' planned extent of testing is measured by budgeted audit hours.

Hypothesis four predicted that the planned extent of tests of details would vary systematically with changes in the volatility of the client's external business environment. However, the main effect for the external business environment (Factor B) was *not* statistically significant ($F < 1.0$). Thus, when based on budgeted audit hours responses, hypothesis four is not supported. On the other hand, one cannot draw the conclusion, from these results, that auditors' budgeted hours do *not* vary systematically with changes in the external business environment. When the null hypothesis is not rejected, one must reserve judgment as to the veracity of the alternative, or research, hypothesis. In this study, it may be that the external environment did indeed systematically influence subjects' budgeted audit hours responses, but the way in which the external environment was operationalized, i.e., as a between-subjects factor, limited the statistical power of the test of the effect of the external environment. This is a particularly viable possibility in this case since subjects varied widely in their average budgeted hours responses. With between-subjects factors, such individual between-subject differences are "superimposed" over whatever treatment effects may have been produced by the experimental manipulation of environmental volatility (Keppel [1982], pg. 369).

Hypothesis five predicts that the effect of the internal control structure on auditors' planned extent of tests of details will depend on conditions in the client's internal operating environment. This is a prediction of an ordinal interaction between the internal operating environment (Factor C) and the internal control structure (Factor D).

However, as revealed in Table 4.1, the C x D interaction was *not* significant. The results of this study *are not consistent with* the hypothesis that the effect of the internal control structure on auditors' judgments concerning the appropriate number of hours to budget for tests of details depends on conditions in the client's internal operating environment. Thus, hypothesis five is not supported when based on budgeted audit hours responses.

Hypothesis six predicts that the effect of the results of analytical procedures on auditors' planned extent of tests of details will depend on the effectiveness of the internal control structure. This is a prediction of an ordinal interaction between the internal control structure (Factor D) and analytical procedures (Factor E). However, as revealed in Table 4.1, the D x E interaction was *not* significant. The results of this study *are not consistent with* the hypothesis that the effect of the results of analytical procedures on auditors' judgments concerning the appropriate number of hours to budget for tests of details depends on the effectiveness of the internal control structure. Thus, hypothesis six is not supported when based on budgeted audit hours responses.

Hypothesis seven predicts that the effect of the results of analytical procedures on auditors' planned extent of tests of details will depend on the degree of volatility in the client's external business environment. This is a prediction of an ordinal interaction between the external business environment (Factor B) and analytical procedures (Factor E). However, as revealed in Table 4.1, the B x E interaction was *not* significant. The results of this study *are not consistent with* the hypothesis that the effect of analytical procedures

on auditors' judgments concerning the appropriate number of hours to budget for tests of details depends on the volatility in the client's external business environment. Thus, hypothesis seven is not supported when based on budgeted audit hours responses.

Joyce [1976] also used budgeted hours as a dependent variable. He found that, when all his subjects were pooled together, there were no statistically significant effects at all. One possible explanation for his result may be that high inter-subject response variability resulted in large within-cell variance terms and low power. Therefore, the current study employed a second measure of subjects' planned extent of testing which was expected to have lower within-cell variability, and thus greater power at detecting any effects which may be present. The next section deals with this issue.

4.2.3 Extent-of-Testing Ratings

Hypotheses one through seven were reevaluated using subjects' extent-of-testing ratings, rather than budgeted audit hours, as the dependent variable in the analyses. Cell means and standard deviations based on extent-of-testing ratings are presented in Table 4.6. As with budgeted audit hours, in all cases, subjects' extent-of-testing ratings were greater when the internal operating environment was unfavorable than when it was favorable, *ceteris paribus*. Similarly, extent-of-testing ratings were greater when internal control was weak than when it was strong, and when the results of analytical procedures were favorable than when they were unfavorable. This pattern of results is consistent with hypotheses one, two, and three. In addition, subjects' extent-of-testing ratings were consistently greater in the volatile external business environment than in the stable environment, *ceteris paribus*.

Table 4.6
Extent-of-Testing Ratings
Cell Means and Standard Deviations

Panel A
Collapsed Across Experience Level

		Internal Operating Environment				
		<u>Unfavorable</u>	<u>Favorable</u>	<u>Unfavorable</u>	<u>Favorable</u>	<u>Averages</u>
		Results of Analytical Procedures	Results of Analytical Procedures	Results of Analytical Procedures	Results of Analytical Procedures	
		<u>Unfavor- able</u>	<u>Favor- able</u>	<u>Unfavor- able</u>	<u>Favor- able</u>	
Strong Internal Control Structure	Stable	13.19	11.22	9.82	6.52	10.19
	External	(1.80)	(2.52)	(2.62)	(2.64)	
	Environment					10.52
	Volatile	13.56	12.07	10.63	7.11	10.84
Weak Internal Control Structure	External	(2.08)	(2.22)	(2.29)	(2.49)	
	Environment					
	Stable	15.11	12.85	12.00	10.56	12.63
	External	(1.74)	(2.20)	(2.22)	(2.42)	
Weak Internal Control Structure	Environment					13.00
	Volatile	15.78	13.85	12.78	11.04	13.36
	External	(1.60)	(2.01)	(1.93)	(2.33)	
	Environment					
<u>Averages</u>		14.41	12.50	11.31	8.81	
		13.46		10.06		11.76

Panel B
Grouped by Experience Level

<u>Students</u>	<u>Seniors</u>	<u>Managers</u>
11.71	11.79	11.77
(3.45)	(2.96)	(3.42)

While this is consistent with hypothesis four, it should be noted that such a pattern was not observed in subjects' budgeted-audit-hours responses. While subjects' extent-of-testing ratings varied systematically with changes in the external business environment, the external business environment had no clear effect on budgeted hours.

The average cell means and standard deviations for students', seniors', and managers' extent-of-testing ratings are also presented in Table 4.6. Students' extent-of-testing ratings are much more consistent with the ratings of seniors and managers than were their responses based on budgeted audit hours. Audit experience is less critical in judging the relative extent of testing for various audit situations than it is in determining the actual number of audit hours that should be performed. Apparently, students learn enough from their course work to allow them to make reasonable judgments concerning the manner in which changes in various audit circumstances might alter the extent of audit testing. However, students generally appear to be very uncertain about the actual number of hours that it might take to complete an audit of accounts receivable under any given set of conditions.

The ANOVA table for subjects' extent-of-testing ratings is presented in Table 4.7. Several differences should be noted between the ANOVA table based on extent-of-testing ratings and the ANOVA table based on budgeted hours. First, there was no main effect for experience level when extent-of-testing ratings were used to measure subjects' planned extent of testing. Average extent-of-testing ratings for students, seniors and managers were nearly identical -- 11.71, 11.79, and 11.77, respectively. However, average budgeted audit hours for students was significantly different from those of seniors and managers. Students'

Table 4.7
Extent-of-Testing Ratings
ANOVA Table

Source	SS	df	MS	F	P	Eta ²
Experience Level (A)	0.48	2	0.238	0.01		.000
Environm'l Volatility (B)	52.08	1	52.083	3.19	.080	.011
A x B	40.63	1	20.313	1.25	.296	.009
<u>S/AB</u>	<u>782.56</u>	<u>48</u>	<u>16.303</u>			
Intnl. Oprtng Envrn. (C)	1247.12	1	1247.120	220.31	.000 **	.270
A x C	51.70	2	25.850	4.57	.015 *	.011
B x C	0.08	1	0.083	0.01		.000
A x B x C	1.12	2	0.562	0.10		.000
<u>C x S/AB</u>	<u>271.72</u>	<u>48</u>	<u>5.661</u>			
Internal Control (D)	665.04	1	665.037	144.73	.000 **	.144
A x D	8.70	2	4.350	0.95		.002
B x D	0.15	1	0.148	0.03		.000
A x B x D	0.81	2	0.405	0.09		.000
<u>D x S/AB</u>	<u>220.56</u>	<u>48</u>	<u>4.595</u>			
C x D	37.93	1	37.926	22.10	.000 **	.008
A x C x D	1.03	2	0.516	0.30		.000
B x C x D	0.59	1	0.593	0.35		.000
A x B x C x D	4.31	2	2.155	1.26	.294	.001
<u>C x D x S/AB</u>	<u>82.39</u>	<u>48</u>	<u>1.716</u>			
Analyt'l Procedures (E)	524.48	1	524.481	105.58	.000 **	.113
A x E	57.42	2	28.711	5.78	.005 **	.012
B x E	0.15	1	0.148	0.03		.000
A x B x E	9.25	2	4.627	0.93		.002
<u>E x S/AB</u>	<u>238.44</u>	<u>48</u>	<u>4.968</u>			
C x E	9.48	1	9.481	5.39	.024 *	.002
A x C x E	0.75	2	0.377	0.21		.000
B x C x E	3.00	1	3.000	1.70	.197	.001
A x B x C x E	1.01	2	0.507	0.29		.000
<u>C x E x S/AB</u>	<u>84.50</u>	<u>48</u>	<u>1.760</u>			
D x E	14.08	1	14.083	6.84	.011 *	.003
A x D x E	4.18	2	2.090	1.01	.370	.001
B x D x E	0.08	1	0.083	0.04		.000
A x B x D x E	6.01	2	3.007	1.46	.242	.001
<u>D x E x S/AB</u>	<u>98.89</u>	<u>48</u>	<u>2.060</u>			
C x D x E	32.23	1	32.231	21.67	.000 **	.007
A x C x D x E	1.14	2	0.572	0.38		.000
B x C x D x E	0.01	1	0.009	0.01		.000
A x B x C x D x E	0.48	2	0.238	0.16		.000
<u>C x D x E x S/AB</u>	<u>71.39</u>	<u>48</u>	<u>1.487</u>			
Total	4625.97	431				

* .01<P<.05

** P<.01

average hours was 166.41 compared to that for seniors of 26.77 and for managers of 39.60.

Hypothesis one predicts that the internal operating environment will influence the extent of tests of details. In order to test hypothesis one based on extent-of-testing ratings, it was necessary to begin by analyzing the experience-level x internal-operating-environment interaction. The nature of this interaction is illustrated in Figure 4.4.

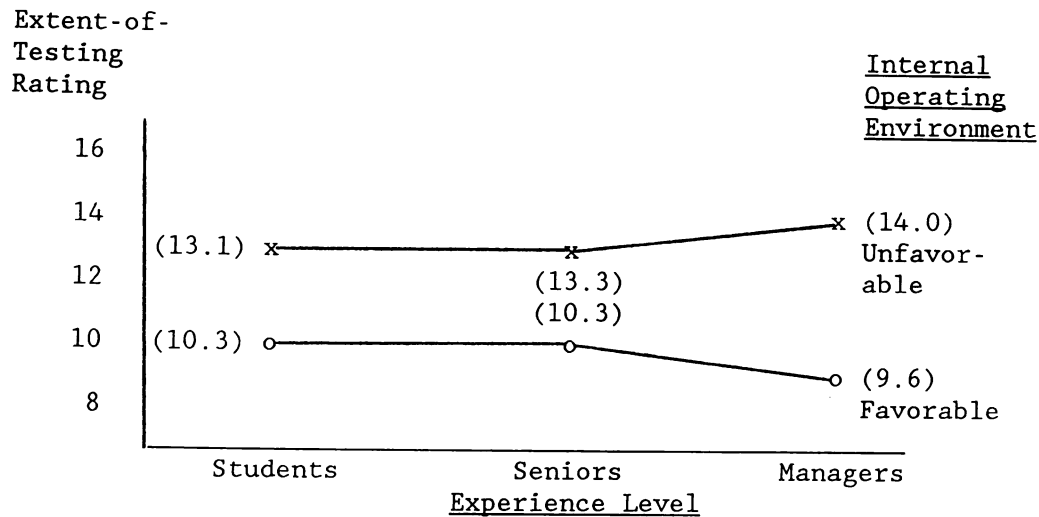


Figure 4.4
Extent-of-Testing Ratings
Interaction Between Experience Level
and Internal Operating Environment (A x C)

While the graph indicates that the largest difference between the favorable and unfavorable internal operating environment conditions was for managers, the ordinal nature of the interaction is apparent in that subjects' responses in the unfavorable condition were consistently higher than in the favorable condition. In addition, the simple effects of internal operating environment at each of the three levels of experience were significant ($P < .001$). These results support

hypothesis one. However, internal operating environment was also involved in a 3-way interaction with internal control and analytical procedures. This 3-way interaction needs to be analyzed before conclusions can be reached regarding hypothesis one. This is done below.

The nature of the 3-way internal-operating-environment x internal-control x analytical-procedure ($C \times D \times E$) interaction is shown in Figure 4.5, Panels A and B. Figure 4.5 shows the 3-way interaction plotted as two simple $C \times D$ interactions at both levels of factor E. (The same patterns hold when the simple $C \times E$ interactions at d_1 and d_2 and the simple $D \times E$ interactions at c_1 and c_2 are plotted.) The results of the analyses of simple $C \times D$ interactions at each of the two levels of E are presented in Table 4.8. As is apparent from Panel A of Figure 4.5, there is no $C \times D$ interaction ($F < 1.0$) when the results of analytical procedures are unfavorable (e_1). However, Panel B of Figure 4.5 illustrates that when the results of analytical procedures are favorable (e_2), there is a significant $C \times D$ interaction ($P < .001$). The simple interactions are not significant at the unfavorable level of the third factor (i.e., $C \times E$ at d_1 ; $D \times E$ at c_1), but are significant when the third factor is favorable. (Again, this same pattern of results holds when the $C \times D \times E$ interaction is decomposed in terms of simple $C \times E$ interactions and simple $D \times E$ interactions.)

To complete the testing of hypothesis one, the simple effects of internal operating environment at the unfavorable level of analytical procedures (e_1), and at both levels of internal control at the favorable level of analytical procedures (d_1e_2 and d_2e_2) were examined. Each simple effect was statistically significant ($P < .001$). A summary of

Extent-of-
Testing
Rating

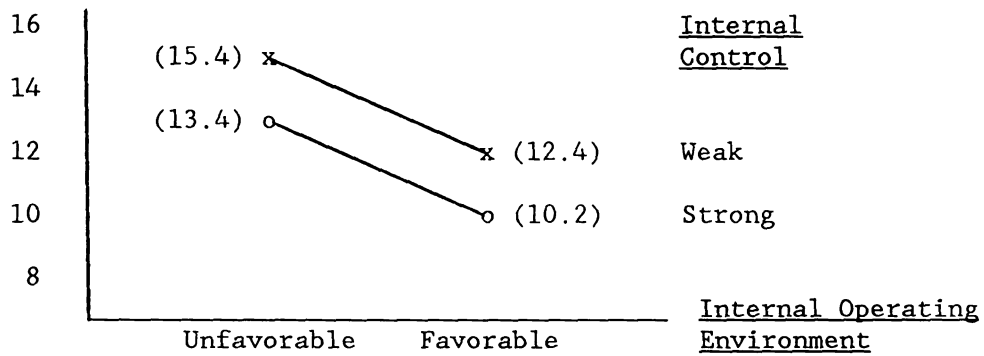


Figure 4.5
Extent-of-Testing Ratings
Three-Way Interaction Between Internal Operating Environment,
Internal Control, and Analytical Procedures (C x D x E)

Panel A
Internal Operating Environment x Internal Control (C x D)
Interaction Given UNFAVORABLE Results of Analytical Procedures (e_1)

Extent-of-
Testing
Rating

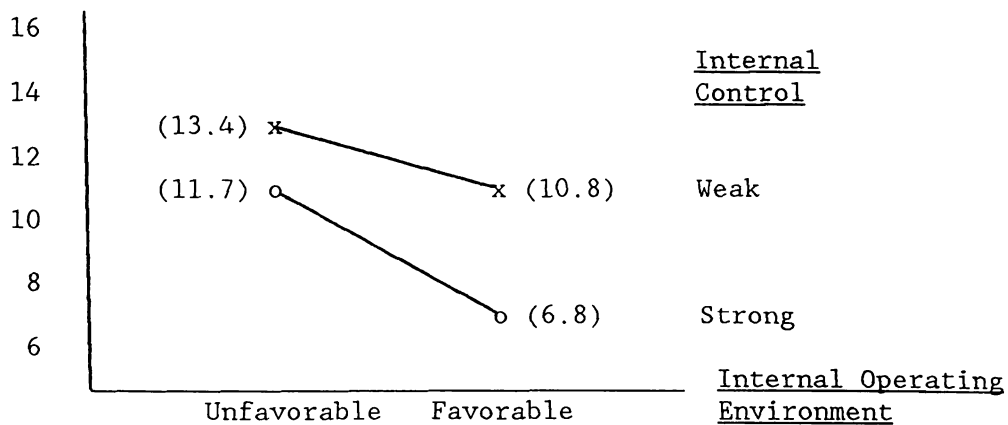


Figure 4.5 (cont'd)

Panel B
Internal Operating Environment x Internal Control (C x D)
Interaction Given FAVORABLE Results of Analytical Procedures (e_2)

Table 4.8
Extent-of-Testing Ratings
Analysis of Simple Interactions

Panel A

Internal Operating Environment x Internal Control (C x D)
Given UNFAVORABLE Results of Analytical Procedures (E)

Source	SS	df	MS	F	P	Eta ²
Intrnl Oprtng Envrn. (C)	519.56	1	519.560	199.99	.000 **	.311
Internal Control (D)	242.78	1	242.782	70.13	.000 **	.146
C x D	0.12	1	0.116	0.07	N/S	.000

Table 4.8 (cont'd.)

Panel B

Internal Operating Environment x Internal Control (C x D)
Given FAVORABLE Results of Analytical Procedures (E)

Source	SS	df	MS	F	P	Eta ²
Intrnl Oprtng Envrn. (C)	737.04	1	737.042	141.43	.000 **	.303
Internal Control (D)	436.34	1	436.338	148.33	.000 **	.179
C x D	70.04	1	70.042	46.87	.000 **	.029

** P<.01

the analysis of the 3-way interaction is presented in Figure 4.6. Panel A of Figure 4.6 summarizes the test of hypothesis one. The significance of each of the simple effects of internal operating environment, combined with the ordinal nature of each of the simple interactions involving internal operating environment, provide support for hypothesis one. As predicted, auditors' judgments of the appropriate extent of audit testing vary systematically with changes in the internal operating environment, with more testing deemed appropriate when conditions in the internal operating environment suggest an increased likelihood of the occurrence of material misstatements in the accounting records. Thus, the conclusion is that auditors, when planning the audit, do indeed take into consideration factors that affect the likelihood of the occurrence of errors or irregularities in the accounting records. This conclusion differs from that of Kaplan and Reckers' [1984] study which found no evidence that auditors' judgments are influenced by the integrity of management - a factor which the auditing literature states should affect the auditor's assessment of the likelihood that errors or irregularities may have occurred. However, Kaplan and Reckers varied only management integrity, whereas the current research varied the aggregate effect of four factors. Further research of the individual and interactive effects of factors which affect the likelihood of misstatements on auditors' judgments is deemed warranted.

Hypothesis two predicts that the internal control structure will influence the extent of tests of details. To test hypothesis two based on extent-of-testing ratings, it was necessary to calculate the simple effects of internal control at the unfavorable level of analytical procedures (D at e_1), and at both levels of internal operating

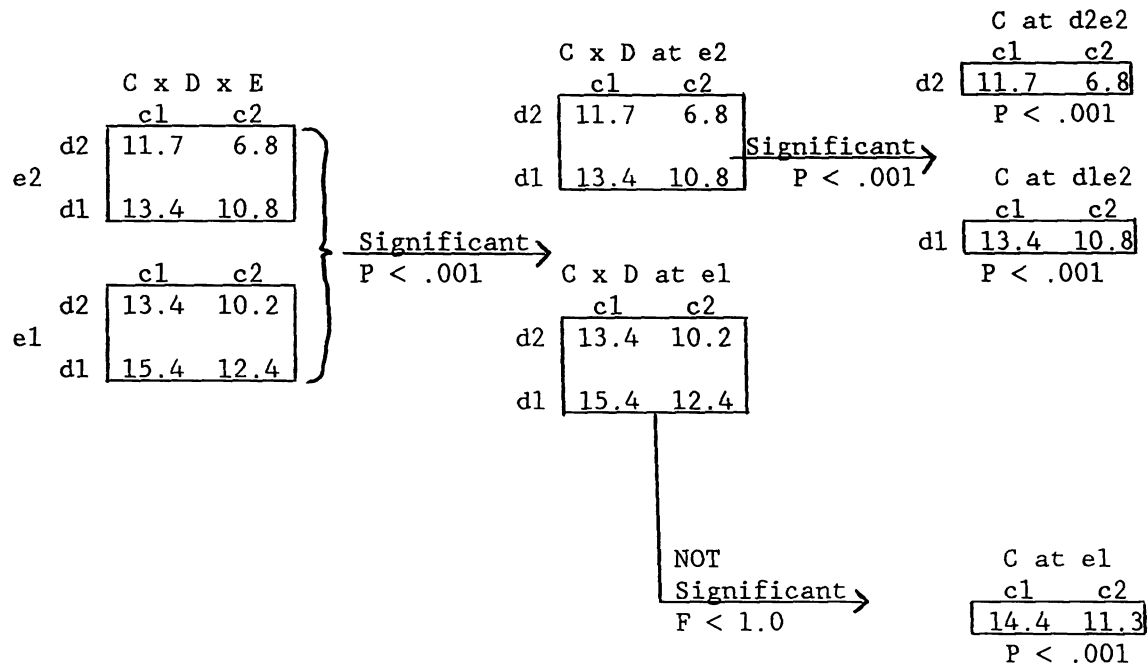


Figure 4.6
Extent-of-Testing Ratings
Analysis of C x D x E Interaction

Panel A
Test of Hypothesis One
C x D at e_1 & e_2

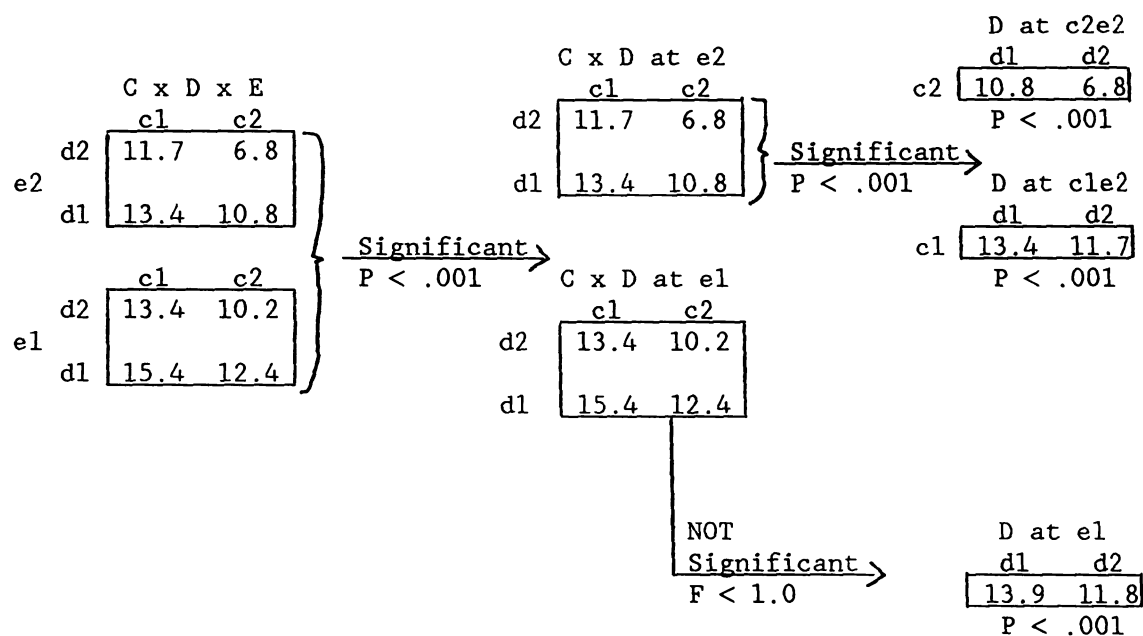


Figure 4.6 (cont.)

Panel B
Test of Hypothesis Two
C x D at e_1 & e_2

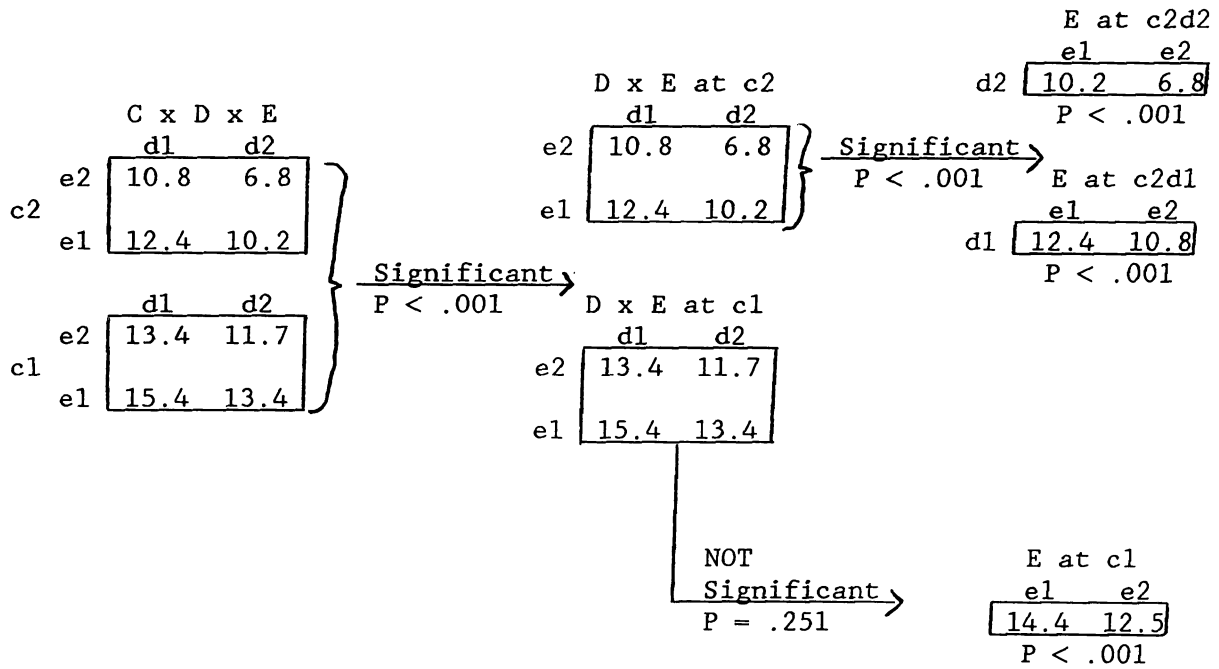


Figure 4.6 (cont.)

Panel C
Test of Hypothesis Three
D x E at c_1 & c_2

environment at the favorable level of analytical procedures (D at c_1e_2 and D at c_2e_2) (see Figure 4.6, Panel B). Each of these simple effects was statistically significant ($P < .001$). Thus, hypothesis two was supported. Therefore, it is concluded that auditors, when planning the appropriate extent of tests of details, do indeed take into consideration the effectiveness of the internal control structure, with more testing deemed appropriate when conditions in the internal control structure suggest an increased likelihood that internal control may not have prevented or detected material misstatements in the accounting records.

Hypothesis three predicts that the results of analytical procedures will influence the extent of tests of details. As in the testing of hypotheses one and two, to test hypothesis three, it was necessary to decompose the C x D x E interaction into simple interactions and then into simple effects. Two simple D x E interactions were formed: D x E at c_1 , and D x E at c_2 . The D x E interaction at c_1 was not statistically significant ($P = .251$), but at c_2 the interaction was significant ($P < .001$). Next, the simple effects of analytical procedures at the unfavorable level of internal operating environment (E at c_1) and at both levels of internal control at the favorable level of internal operating environment (E at c_2d_1 and E at c_2d_2) were calculated. All of these simple effects were significant ($P < .001$), supporting hypothesis three. Therefore, it is concluded that auditors, when planning the appropriate extent of tests of details, do indeed take into consideration the favorableness of the results of analytical procedures, with more tests of details of balances deemed appropriate when analytical procedures suggest an increased likelihood

of the presence of material misstatements in the accounting records, than when analytical procedures suggest a lower likelihood of misstatements. A summary of the test of hypothesis three is presented in Figure 4.6, Panel C.

Hypothesis four predicts that the client's external business environment will influence the extent of tests of details. Since the external business environment (factor B) was not involved in any significant interactions, hypothesis four was tested by examining the main effect for external environment. This main effect was statistically significant at an alpha level of .08. Recall that when hypothesis four was testing using budgeted audit hours as the measure of auditors' extent-of-testing judgments, the effect of the external business environment was not significant ($F < 1.0$). One purpose of asking subjects to provide a graphical extent-of-testing rating was to obtain a measure of their extent-of-testing judgments which tends to have less inter-subject response variability than is normally present in budgeted audit hours responses. Since the within-cell variance terms are smaller when using the extent-of-testing ratings, the power of the statistical tests is increased. Given the level of significance ($P=.08$) of the main effect of the external business environment, hypothesis four is tentatively supported. It is therefore concluded that auditors, when planning the appropriate extent of tests of details, do indeed take into consideration the degree of volatility in the client's external business environment, with more tests of details of balances deemed appropriate when conditions in the environment indicate an increase in the likelihood that the client may experience financial trouble in the near future. Such a situation increases the likelihood that the auditor may

be required to defend the adequacy of the audit in court. This, in turn, influences the auditor's acceptable audit risk, and thus, the extent of audit testing. While these findings differ from those of Kaplan's [1985] study which did *not* conclude that there is a significant relationship between the extent of testing and environmental volatility, they are consistent with guidelines in the auditing literature which suggest that the client's business environment should be considered by the auditor when planning the examination. However, given the marginal level of significance ($P = .08$) of the main effect for environmental volatility, caution must be exercised when forming conclusions regarding the effects of environmental volatility on audit planning.

Hypothesis five predicts that the internal operating environment will influence the reliance auditors place on internal control. Specifically, it is predicted that a change in the effectiveness of the internal control structure will have a larger effect on auditors' extent-of-testing judgments when the internal operating environment was unfavorable; i.e., when there is a relatively high likelihood that material errors or irregularities have occurred during the period under audit. This is a prediction of a specific type of C x D interaction, the nature of which is illustrated in Figure 4.7.

The significant 3-way C x D x E interaction required the C x D interaction to be analyzed at both levels of factor E. Panel A of Figure 4.5 revealed that the C x D interaction was not significant ($F < 1.0$) when the results of analytical procedures suggest there is a high likelihood that errors are present in the pre-audit financial statements (e_1). Even more interesting, however, is the nature of the C x D interaction when the results of analytical procedures are favorable.

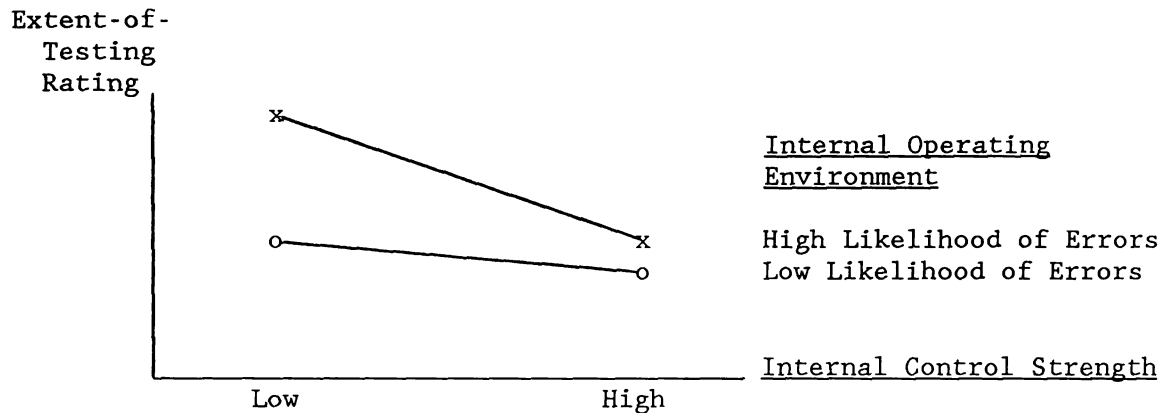


Figure 4.7
Hypothesized Interaction Between
Internal Operating Environment (C) and Internal Control (D)

The form of the C x D interaction at e_2 was revealed in Panel B of Figure 4.5. The hypothesized interaction was based on the critical combination of high likelihood of errors and poor internal controls. Under these conditions, it was expected that the auditor would require extensive testing. Under any of the other three possible combinations of conditions (i.e., high likelihood of errors & good internal controls, low likelihood of errors & poor internal controls, low likelihood of errors & good internal controls) the auditor was expected to require a relatively small amount of testing since good internal controls would compensate for high likelihood of errors, and low likelihood of errors would not require a strong set of internal controls. However, the observed interaction in Panel B of Figure 4.5 reveals that the critical combination of conditions was low likelihood of errors & strong internal controls. Auditors apparently do not feel that strong controls will compensate adequately for an internal environment which results in a high likelihood of errors, or that an internal environment which is conducive to a low likelihood of errors compensates for weak internal

controls. It should be noted that the likelihood-of-errors x internal-control-strength interaction observed by Libby, Artman, and Willingham [1985] (LAW) indicated that a change in control strength has its greatest effect on auditors' control-reliance judgments when there is a *high* likelihood of errors. It would seem to follow that the same set of conditions which produce a large change in auditors' control-reliance judgments should also produce a large change in the extent of tests of details. For instance, a large decrease in control reliance would produce a larger increase in the extent of tests of details than would a small decrease in control reliance. However, the results of the current study indicate that LAW's results *cannot* be extended to auditors' extent-of-testing judgments. In the context of extent-of-testing judgments, a change in control strength has its greatest effect when the likelihood of errors is *low*, rather than high. The observed C x D interaction is thus opposite to the particular C x D interaction predicted by hypothesis five. However, the observed interaction is consistent with the audit risk model.

It is important to keep in mind that when analytical procedures suggest a high likelihood of errors in the pre-audit financial statements, the low likelihood of errors and strong internal controls combination does not produce a critical combination. Thus, there is *no* interaction between the internal operating environment and internal control when the results of analytical procedures are unfavorable. Apparently, all three conditions must be favorable before a critical combination of evidence is obtained. Specifically, not only must there be a low likelihood of errors and strong internal controls, but the results of analytical procedures must also be favorable before a

significant interaction between the different types of evidence is obtained.

In summary, it is concluded that when any of the three types of evidence above suggest a *high* likelihood of misstatements, then the effects of other evidence on auditors' extent-of-detailed-testing judgments are attenuated. Only when all three types of evidence are favorable is the auditor willing to reduce the extent of testing to a great degree. This is analogous to Ashton's [1974] observation that when auditors discover a weakness in internal controls, the results of other tests of the control structure seem to have little effect on the overall evaluation of the structure. Furthermore, these results are consistent with evidence in the psychology literature which has found that negative information tends to be weighted heavier than positive information (Anderson & Alexander [1971], Hamilton & Zanna [1972], and Hodges [1974]). Finally, these observed results are consistent with the audit risk model³. The findings of this research thus lend empirical support to the multiplicative nature of the audit risk model.

Hypothesis six predicts that the effectiveness of the internal control structure will influence the extent to which the auditor is influenced by the results of analytical procedures when planning the extent of audit testing. This is a prediction of an interaction between internal control and analytical procedures (D x E). This interaction was found to be significant ($P = .011$). However, as with the test of hypothesis five, the significance of the 3-way C x D x E interaction meant that simple interactions must be examined in order to test the hypothesis. Accordingly, the D x E interaction was analyzed at both levels of the third factor -- factor C. Panel A of Figure 4.8 reveals

the D x E interaction to be not significant ($P = .251$) when the internal operating environment is unfavorable (c_1). This suggests that when conditions in the internal operating environment are such that there is a relatively high likelihood that material misstatements may have occurred during the period under audit, the amount of audit assurance that auditors are willing to place on the results of analytical procedures is not influenced by the effectiveness of the internal control structure. However, this is not the case when conditions in the internal operating environment are such that there is a low likelihood that misstatements have occurred. As revealed by Panel B of Figure 4.8, when internal operating environment is strong, there is a significant interaction between internal control and analytical procedures ($P < .001$). Apparently, when conditions in the internal operating environment are such that there is a low likelihood that material errors or irregularities have occurred during the period under audit, the amount of reliance that auditors are willing to place on analytical procedures depends on the effectiveness of the internal control structure, as hypothesized. When controls are strong and when conditions in the internal operating environment are favorable, then an increase in the favorableness of analytical procedures results in the auditor being willing to reduce his/her extent of testing by a greater amount than when either controls or the internal operating environment, or both, are unfavorable. It can also be stated that when the results of analytical procedures are favorable and when conditions in the internal operating environment are favorable, then an increase in the strength of the internal control structure will result in the auditor being willing to reduce his/her extent of testing by a greater amount

Extent of
Testing
Rating

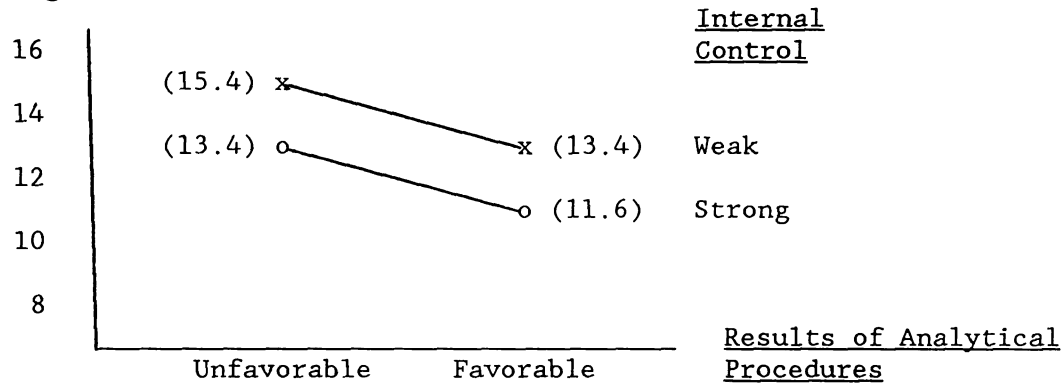


Figure 4.8

Extent-of-Testing Ratings

Three-Way Interaction Between Internal Operating Environment,
Internal Control, and Analytical Procedures (C x D x E)

Panel A

Internal Control x Analytical Procedures (D x E)

Interaction Given UNFAVORABLE Internal Operating Environment (c_1)

Extent of
Testing
Rating

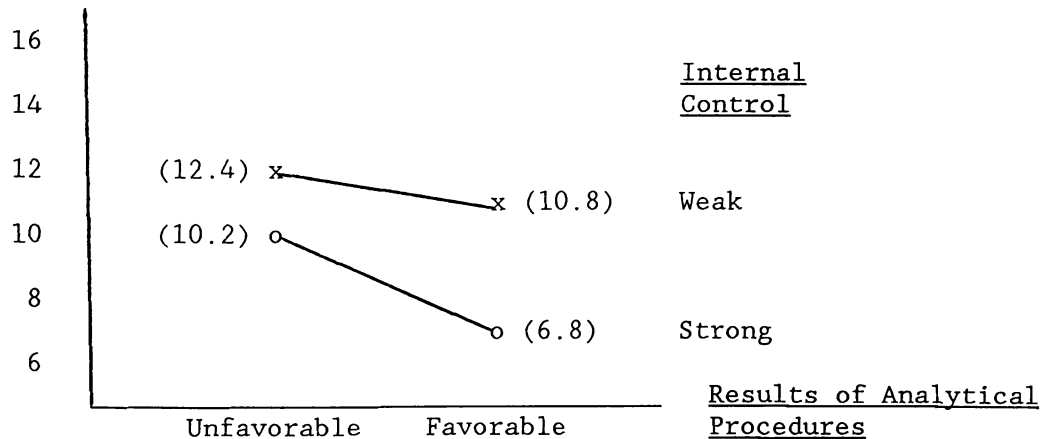


Figure 4.8 (cont'd)

Panel B

Internal Control x Analytical Procedures (D x E)

Interaction Given FAVORABLE Internal Operating Environment (c_2)

than when either the results of analytical procedures or conditions in the internal operating environment, or both, are unfavorable. Thus, the veracity of hypothesis six is conditional upon the state of the internal operating environment. When the internal environment is favorable, the results are consistent with hypothesis six. However, when the internal environment is unfavorable, hypothesis six is not supported. The conclusion is that when any of the three sources of audit information examined in this research - the internal operating environment, the internal control structure, or the results of analytical procedures - suggest a *high* likelihood of misstatements, then the effects of the other sources of information on auditors' extent-of-detailed-testing judgments are reduced. Again, this is consistent with Ashton's [1974] findings, with evidence from the psychology literature (Anderson & Alexander [1971], Hamilton & Zanna [1972], and Hodges [1974]), and is also consistent with the multiplicative nature of the audit risk model.

Hypothesis seven predicts an interaction between the external business environment and analytical procedures (B x E). As was indicated in Table 4.7, this interaction was not significant ($F < 1.0$). Thus, hypothesis seven was not supported. However, SAS 56 suggests that the effects of analytical procedures on auditors' extent-of-tests-of-details judgments *should* depend on the volatility of the client's business environment.

It is important for the auditor to understand the reasons that make relationships plausible because data sometimes appear to be related when they are not, which could lead the auditor to erroneous conclusions ... As higher levels of assurance are desired from analytical procedures, more predictable relationships are required to develop the expectation. Relationships in a stable environment are usually more predictable than relationships in a dynamic or unstable environment. (AICPA [1988]).

This study has found no evidence to support the suggestion in SAS 56 that the extent to which auditors rely on analytical procedures depends on the stability of the external business environment. It is possible that auditors may not be considering the volatility in the environment - or the predictability of expected relationships - when relying on analytical procedures. This issue should be explored further in future research.

Finally, it should be noted that in Table 4.7 that there was a significant internal-operating-environment x analytical-procedures (C x E) interaction ($P = .024$). Since these two factors were involved in the 3-way C x D x E interaction, the simple C x E interactions at both levels of factor D were examined. The results were nearly identical to those in the analysis of the D x E and the C x D interactions. When internal control is weak, the amount of audit assurance that auditors are willing to place on the results of analytical procedures is not influenced by conditions in the internal operating environment; i.e., the internal-operating-environment x analytical-procedures interaction is not significant ($P = .142$). However, when internal control is strong, there is a significant interaction between the internal operating environment and analytical procedures ($P < .001$). A change in the favorableness of the results of analytical procedures will have its greatest effect on auditors' extent-of-testing judgments when the other two factors (internal control and the internal operating environment) are both favorable. Similarly, a change in the favorableness of conditions in the internal operating environment will have its greatest effect on auditors' judgments when internal control and analytical procedures are favorable. As stated previously, it appears that when

any of the three types of evidence suggest a *high* likelihood of misstatements, the effects of other evidence on auditors' extent-of-detailed-testing judgments are attenuated. Only when all three types of evidence are favorable is the auditor willing to reduce the extent of testing to a great degree. Again, this is consistent with the psychology literature and with the audit risk model.

4.2.4 Weighting of Evidence

In order to determine the relative importance, or weight, of each factor and interaction of factors in subjects' judgments, the eta-squared index (n^2) was calculated for each factor and interaction. Eta-squared corresponds to the squared multiple-correlation coefficient (R^2) in the multiple-regression literature and provides a relative measure of the importance of each factor. The eta-squared indices based on all subjects' judgments were presented in Table 4.7. The factor with the largest eta-squared index was the internal operating environment ($n^2 = .270$). Next in importance was internal control ($n^2 = .144$), followed by analytical procedures ($n^2 = .113$) and external environmental volatility ($n^2 = .011$). The internal operating environment apparently was weighted heavier than any other factor in auditors' judgments of the appropriate extent of tests of details. Internal control and analytical procedures each received roughly half as much weight as the internal operating environment, but were nearly equal to one another in importance and were both significant factors in auditors' judgments. The external operating environment was less important, accounting for just over one percent of the variance in subjects' extent-of-testing ratings.

To examine whether subjects with different levels of audit experience weighed factors differently, eta-squared indices were

calculated separately for students, seniors, and managers. These results are presented in Table 4.9.

Table 4.9
Extent-of-Testing Ratings
Eta-Squared Indices
Students, Seniors, and Managers

	<u>Eta-Squared</u>		
	<u>Students</u>	<u>Seniors</u>	<u>Managers</u>
Environmental Volatility (B)	.065	.002	.000
Internal Operating Environment (C)	.162	.267	.412
Internal Control Procedures (D)	.121	.209	.097
Analytical Procedures (E)	.182	.129	.042
Sum of Two-Way Interactions	.012	.010	.020

Table 4.9 reveals some interesting findings. First, the importance of environmental volatility declines as experience increases. This observation was supported by the F-ratios in three ANOVAs based on separate extent-of-testing ratings of students, seniors, and managers. The ANOVA based on the ratings of students alone reveals environmental volatility to be a statistically significant factor in the judgments of students ($P = .046$). However, the ANOVAs for seniors and managers showed environment volatility to be insignificant to those two groups of subjects ($F < 1.0$ for both groups). Apparently, auditors do not feel that changes in the volatility of the client's business environment are matters which should affect the extent of tests of details.

Second, while the internal operating environment was significant for all three groups of subjects, its relative importance increases with experience. The internal operating environment is, by far, the most important factor in managers' judgments. The eta-squared index for the internal operating environment for managers was .412, while the eta-

squared index for the second-most important factor to managers, internal control procedures, was considerably lower at .097. The internal operating environment is also the single most important factor for seniors. However, its eta-squared index, .267, was just over half that of the corresponding index for managers, and the eta-squared index for seniors' second-most important factor, internal control procedures, was only moderately lower than .267 at .209.

Third, Table 4.9 reveals that the relative importance of analytical procedures declines as experience increases. Analytical procedures is the most important factor to students ($n^2 = .182$), but is a much less important factor to seniors and managers ($n^2 = .129$ and .042, respectively). Managers, in particular, apparently do not feel that analytical procedures are of much importance in the determination of the appropriate extent of tests of details. This is curious since SAS 58 has recently reemphasized the importance of analytical procedures in audit planning. Perhaps the current members of the Auditing Standards Board believe that analytical procedures are of more importance in audit planning than do audit managers in general.

Finally, the sum of the eta-squared indices for all two-way interactions of factors was very low across all three subject groups, ranging from .01 for seniors to .02 for managers. Although statistically significant interactive effects were observed, the magnitude of these effects was small when compared to the factors' main effects. This suggests that a relatively small degree of configurality is present in auditors' audit planning judgment processes.

4.3 EFFECTS OF EXPERIENCE

4.3.1 Judgmental Agreement Based on Correlations

To examine the effects of experience on audit judgment and judgmental agreement between auditors, subjects were asked to perform two audit tasks -- plan the extent of audit testing for accounts receivable, and evaluate the internal control structure over receivables. Hypothesis eight predicted there would be a larger difference between the levels of judgmental agreement of experienced and inexperienced auditors when planning the extent of testing than when evaluating internal control. The mean Pearson product-moment correlation coefficient between subjects' pair-wise judgments was initially used as the measure of judgmental agreement between auditors. The mean correlations across experience level and audit task are presented in Table 4.10.

Table 4.10
Inter-Auditor Consensus as Measured by
Pair-Wise Correlation Coefficients

		<u>Task</u>	
		<u>Audit</u>	<u>Internal Control</u>
		<u>Planning</u>	<u>Evaluation</u>
<u>Experience</u> <u>Level</u>	<u>Practitioners</u>	0.737 (0.163)	0.702 (0.243)
	<u>Students</u>	0.724 (0.195)	0.742 (0.169)

Standard deviations are shown in parentheses.

The mean correlation between pair-wise extent-of-testing ratings for practitioners was .737 with a standard deviation of .163 and a range of .145 to .992. For students, the mean correlation was .724 with a standard deviation of .195 and a range of -.020 to .982. These results

reveal that the practitioners had a higher level of agreement among themselves, as measured by pair-wise correlations, concerning the appropriate extent of audit testing than did the students. However, the magnitude of the difference is not statistically significant ($F < 1.0$).

The mean correlation between pair-wise internal control evaluations of practitioners was .702 with a standard deviation of .243 and a range of -.231 to .959. For students, the mean correlation was unexpectedly higher than that of practitioners -- .742 with a standard deviation of .169 and a range from .298 to .987. However, as with the extent-of-testing ratings, the difference between practitioners' mean correlation and that of students was not statistically significant ($P = .196$).

Hypothesis eight was tested via a 2 x 2 ANOVA in which the Pearson product-moment correlation between pair-wise judgments for students and practitioners was the dependent variable, and the independent variables were experience level and audit task. The ANOVA results are presented in Table 4.11. The experience-level x audit-task interaction was not significant ($P = .182$). This indicates that the difference between the level of judgmental agreement - when measured by the correlation coefficient - among practitioners and the level of agreement among auditing students when determining the appropriate extent of tests of details of balances is no greater than the corresponding difference when evaluating the internal control structure. This finding is not consistent with hypothesis eight. Two explanations are offered for this outcome. First, it is possible that the role of professional experience is no more important in audit planning than it is in internal control evaluation. However, this explanation does not seem likely since

previous research suggests that internal control evaluation is a relatively structured task at which students are normally able to do quite well (e.g., Ashton and Kramer [1980]) while audit planning is a more difficult, less structured task (Joyce [1976]) at which students should be less adept. Second, it is possible that the use of the correlation coefficient as the measure of judgmental agreement masked some of the differences in the judgments of the two groups of subjects and is responsible for the lack of significance in the experience-level x audit-task interaction. These two possibilities are addressed in the next section.

Table 4.11
Inter-Auditor Consensus Based on Correlation Coefficients
Students & Practitioners
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	0.02	1	0.018	0.43	
S/A (Error)	7.30	178	0.041		
Audit Task (B)	0.01	1	0.006	0.18	
A x B	0.06	1	0.063	1.79	.182
B x S/A (Error)	6.28	178	0.035		

4.3.2. Judgmental Agreement Based on Mean Absolute Differences

Gaumnitz, Nunamaker, Surdick, and Thomas [1982] have pointed out some of the weaknesses associated with the use of correlation coefficients to measure judgmental consensus among auditors. For instance, while correlations measure the extent to which high (and low) judgments of one subject (relative to that subject's average judgment) tend to be associated with high (and low) judgments of another subject, they tell us nothing about how different the subjects' average judgments are. This weakness appears to be quite serious in the current study.

For example, Student #1's average budgeted audit hours was 36.25 hours with a variance of 262.5. Student #5's average budgeted audit hours was 656.25 with a variance of 42,455.36. Although student #5 felt that each condition required many more hours of audit testing than did student #1, the correlation between their audit judgments was .96.

Wright [1976] advocated the "use of a deviation measure which is not an association metric such as the mean absolute (or squared) difference for pair-wise judgments". Accordingly, the mean absolute differences (MADs) between subjects' judgments were calculated across experience level and audit task. The mean absolute differences between subjects' judgments (CNS_{MAD}) are presented in Table 4.12.

Table 4.12
Inter-Auditor Consensus as Measured by
Mean Absolute Differences (CNS_{MAD})

		<u>Task</u>	
		<u>Audit</u>	<u>Internal Control</u>
		<u>Planning</u>	<u>Evaluation</u>
<u>Experience</u> <u>Level</u>	Practitioners	2.433 (0.843)	2.690 (1.171)
	Students	2.863 (1.285)	2.396 (0.852)

Standard deviations are shown in parentheses.

The CNS_{MAD} between pair-wise extent-of-testing ratings for practitioners was 2.433 with a standard deviation of 0.843 and a range of 1.000 to 4.875. For students, the CNS_{MAD} was 2.863 with a standard deviation of 1.285 and a range of 0.500 to 7.125. These results reveal that the practitioners had a higher level of agreement among themselves, as measured by mean absolute differences between pair-wise judgments, concerning the appropriate extent of audit testing than did the

students. Unlike the results based on correlations, the difference between practitioners' CNS_{MAD} and students' CNS_{MAD} for extent-of-testing ratings ($2.863 - 2.433 = 0.430$) was statistically significant ($P = .008$). This suggests that the use of the correlation coefficient as a measure of judgmental agreement may not be as effective as other possible measures of consensus. Apparently, the correlation coefficient masked some of the differences in the judgments of the two groups of subjects and was responsible for the previously observed lack of significance in the experience-level x audit-task interaction.

The CNS_{MAD} between pair-wise internal control evaluations of practitioners was 2.690 with a standard deviation of 1.171 and a range of 0.625 to 6.375. For students, the CNS_{MAD} was unexpectedly lower than that of practitioners -- 2.396. The standard deviation of students' internal control evaluations was also lower at 0.852 than that of practitioners. The range of students' CNS_{MAD} was 1.000 to 4.875. However, the difference between students' CNS_{MAD} and practitioners' CNS_{MAD} for internal control evaluations ($2.690 - 2.396 = .294$) was not significant ($P = .06$). The pattern of results observed for internal control evaluations was reversed from that observed for extent-of-testing judgments. Apparently, audit experience plays an important role in reducing inter-auditor disagreement in audit planning, but has a less important role in the evaluation of controls. While practitioners showed a significantly higher level of judgmental agreement among themselves than did the students when planning an audit, there was not a significant difference in the consensus levels of practitioners and students when evaluating controls.

A 2 x 2 ANOVA with CNS_{MAD} as the dependent variable and experience

level and audit task as the factors was performed as a test of hypothesis eight. Recall that hypothesis eight predicted an experience-level x audit-task interaction. The results of the ANOVA are presented in Table 4.13.

Table 4.13
Inter-Auditor Consensus Based on
Mean Absolute Differences (CNS_{MAD})
Students & Practitioners
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	0.41	1	0.408	0.35	
S/A (Error)	210.10	178	1.180		
Audit Task (B)	0.99	1	0.990	0.94	
A x B	11.78	1	11.781	11.24	.000 **
B x S/A (Error)	186.61	178	1.048		

** $P < .01$

The experience-level x audit-task interaction was significant ($P < .001$), and is plotted in Figure 4.9.

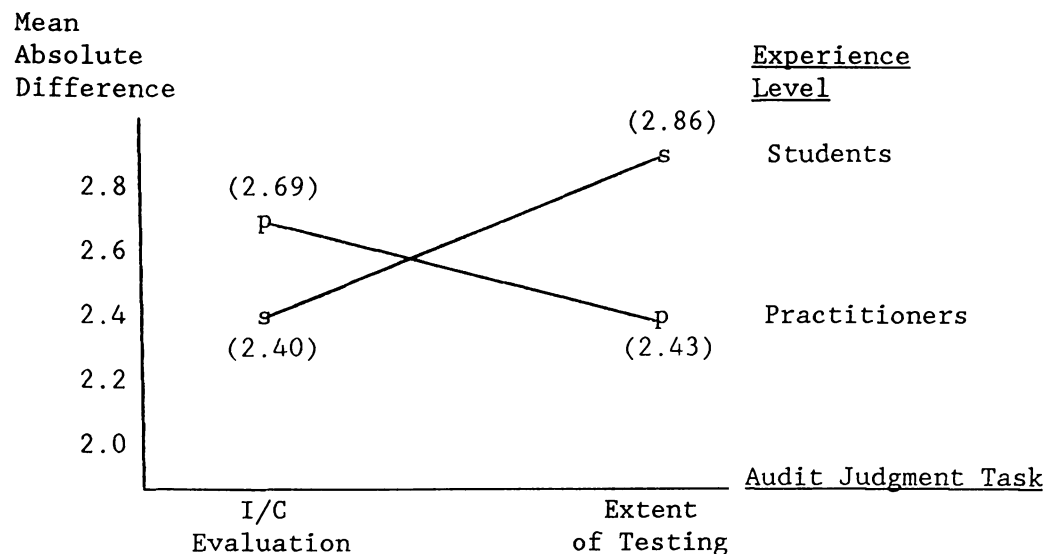


Figure 4.9
Experience Level x Audit Task Interaction (CNS_{MAD})
Students and Practitioners

The significant experience-level x audit-task interaction supports hypothesis eight. Practitioners showed more consensus than did students in the audit planning task, and there was a greater difference between practitioners' and students' consensus levels in the planning task than in the internal control evaluation task. Interestingly, students showed greater consensus than did practitioners in the internal control evaluation task, although this difference was not statistically significant.

The objective of this phase of the research was to examine the effects of knowledge obtained through experience on audit planning judgments. The observed results are consistent with the hypothesized effect. Consistent with previous research (e.g., Ashton and Kramer [1980], audit experience appears to have little effect on internal control evaluations. Apparently, auditing students develop sufficient skills through their course work to enable them to evaluate the relative strength of various internal control structures reasonably well. However, in the area of audit planning, audit experience appears to play a more vital role. Although students learn in their course work how various factors such as internal control effectiveness and the results of analytical procedures might affect the extent of tests of details, this task is less structured and more difficult than internal control evaluation. While students may be acceptable surrogates for practitioners in internal control evaluation tasks, the results of this research suggest that students are not good surrogates for practitioners in less structured tasks such as audit planning.

4.3.3 Decomposition of Consensus

The social judgment version of the lens model describes consensus as a function of cue weighting agreement, judgmental consistency, and agreement on configurality (Hammond, et al. [1975]; Bonner [1988]):

$$\text{Consensus} = f(\text{cue weighting agreement, consistency, configurality})$$

In order to gain a better understanding of the sources of differences in the judgment processes of auditors at differing levels of experience, each of the above elements of consensus was examined. Cue weighting agreement was measured in terms of the mean absolute difference between the predictions of linear regression models of individuals 1 and 2's judgments (G_{MAD}). Judgment consistency was measured as the mean absolute difference between individual i 's judgments and the judgments predicted by the regression model of individual i 's judgments (R_{MAD}). Agreement on configural cue utilization was measured as the mean absolute difference between the residual variances which are unexplained by the linear regression models of individuals 1 and 2 (C_{MAD}). Values for G_{MAD} , R_{MAD} and C_{MAD} across experience level and audit task are presented in Table 4.14. From Table 4.14, it appears as though much of the difference in consensus levels, CNS_{MAD} , is due to differences in agreement on cue weighting, G_{MAD} , with R_{MAD} and C_{MAD} having little influence. To examine this further, three 2×2 ANOVAs were performed with G_{MAD} , R_{MAD} , and C_{MAD} as the dependent variables, and experience level and audit task as the factors. Results of the ANOVAs are presented in Tables 4.15, 4.16, and 4.17. As shown in Table 4.15, the experience-level $\times$ audit-task interaction was significant ($P < .001$) for cue weighting agreement (G_{MAD}), indicating that the extent to which students and practitioners differ in cue weighting agreement depends on the audit

Table 4.14
Mean Values for G_{MAD} , R_{MAD} , and C_{MAD}

	G_{MAD}	R_{MAD}	C_{MAD}
Practitioners -- Audit Planning	2.167 (0.882)	0.766 (0.280)	1.059 (0.424)
Students -- Audit Planning	2.669 (1.361)	0.822 (0.269)	1.054 (0.381)
Practitioners -- I/C Evaluation	2.615 (1.199)	0.722 (0.411)	1.047 (0.484)
Students -- I/C Evaluation	2.096 (0.880)	0.869 (0.526)	1.156 (0.500)

Standard deviations are shown in parentheses.

Table 4.15
Cue Weighting Agreement (G_{MAD})
Students & Practitioners
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	0.01	1	0.007	0.01	
S/A (Error)	219.76	178	1.235		
Audit Task (B)	0.36	1	0.356	0.30	
A x B	23.41	1	23.414	19.74	.000 **
B x S/A (Error)	211.15	178	1.186		

** $P < .01$

task being performed. The nature of this interaction is plotted in Figure 4.10. The pattern of this interaction is nearly identical to that observed for CNS_{MAD} . Specifically, practitioners had a higher level of agreement among themselves regarding the relative weighting of factors in the audit planning task than did students, whereas students agreed more among themselves regarding the relative weighting of factors in the internal control evaluation task than did practitioners.

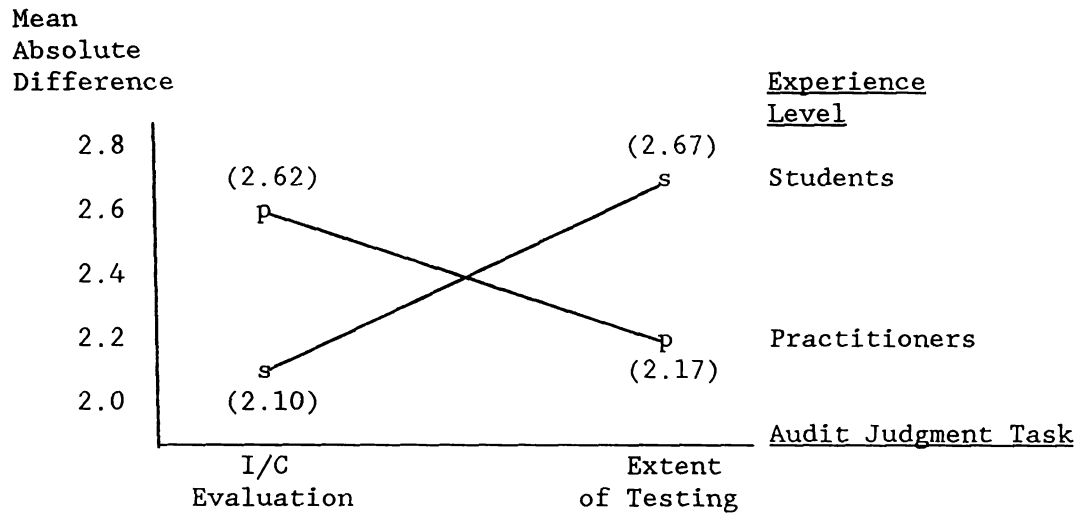


Figure 4.10
Experience Level x Audit Task Interaction (G_{MAD})
Students and Practitioners

Table 4.16
Judgment Consistency (R_{MAD})
Students & Practitioners
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	0.21	1	0.206	1.18	.283
S/A (Error)	6.64	38	0.175		
Audit Task (B)	0.00	1	0.001	0.00	
A x B	0.04	1	0.041	0.33	
B x S/A (Error)	4.69	38	0.123		

Table 4.16 presents the ANOVA for judgment consistency (R_{MAD}). There were no significant effects for R_{MAD} . The experience-level x audit-task interaction for CNS_{MAD} could not have been caused by the student-practitioner difference in R_{MAD} since the difference between practitioners' R_{MAD} 's and students' R_{MAD} 's was over twice as large in internal control evaluation than in audit planning, although this was not large enough to produce an interaction. Thus, differences in the consistency with which auditors at different levels of experience make

judgments is not responsible for the experience-level x audit-task interaction in consensus.

The results of the ANOVA for configural cue utilization agreement (C_{MAD}) are presented in Table 4.17. As with R_{MAD} , there were no significant effects for C_{MAD} . The experience-level x audit-task interaction for $C_{NS_{MAD}}$ could not have been caused by the student-practitioner difference in C_{MAD} since the difference between practitioners' C_{MAD} 's and students' C_{MAD} 's was larger in internal control evaluation than in audit planning, but not large enough to produce an interaction.

Table 4.17
Configural Cue Utilization Agreement (C_{MAD})
Students & Practitioners
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	0.25	1	0.247	1.12	.291
S/A (Error)	39.39	178	0.221		
Audit Task (B)	0.18	1	0.181	0.98	
A x B	0.30	1	0.295	1.61	.206
B x S/A (Error)	32.65	178	0.183		

Differences in the extent to which auditors with differing levels of audit experience employ configurality in their judgment processes is not responsible for the experience-level x audit-task interaction in consensus. Therefore, the significant experience-level x audit-task interaction in auditor consensus ($C_{NS_{MAD}}$), in which the difference between students' and practitioners' audit planning consensus (0.430) is greater than the difference between students' and practitioners' internal control evaluation consensus (-0.294), is apparently due primarily to differences in cue weighting agreement (C_{MAD}).

4.3.4 Decomposition of Experience Levels

In order to better understand the effects of experience and how these effects differ across experience levels, the "practitioners" level of experience was partitioned into two groups -- seniors and managers. Comparisons were then made of the level of judgmental agreement among students with that among seniors, and the level of judgmental agreement of seniors with that of managers. This examination was conducted by first calculating separately the mean absolute differences in the judgments of managers and also in the judgments of seniors. These CNS_{MAD} 's are presented in Table 4.18 along with the previously calculated CNS_{MAD} 's for students.

Table 4.18
Inter-Auditor Consensus as Measured by
Mean Absolute Differences (CNS_{MAD})
Managers, Seniors, & Students

		Task	
		<u>Audit Planning</u>	<u>Internal Control Evaluation</u>
<u>Experience Level</u>	Manager	2.660 (0.858)	2.627 (1.242)
	Senior	2.121 (0.648)	2.769 (1.027)
	Student	2.863 (1.285)	2.396 (0.852)

Standard deviations are shown in parentheses.

The mean CNS_{MAD} for managers' extent-of-testing judgments was 2.660 with a standard deviation of 0.858 and a range of 1.000 to 4.875. For seniors, the mean CNS_{MAD} for extent-of-testing judgments was 2.121 with a standard deviation of 0.648 and a range of 1.000 to 3.875. Comparing the mean CNS_{MAD} of seniors (2.121) with that of students

(2.863) reveals that seniors agree more among themselves than do students concerning the appropriate extent of audit testing. The difference in students' and seniors' CNS_{MAD} 's (0.742) is statistically significant ($P < .001$). This difference is as expected, given that students learn relatively little in their course work concerning the appropriate extent of audit testing under various circumstances, while audit seniors generally have been involved to some extent in audit planning. A further increase in consensus was also expected as one moves from the senior level to the manager level, but such an increase was not observed. A comparison of the mean CNS_{MAD} for seniors (2.136) with that for managers (2.667) indicates that managers agree less among themselves concerning the appropriate extent of testing than do seniors⁴. The magnitude of this difference (0.531) is statistically significant ($P < .001$). While this finding was unexpected, the literature reveals other instances in which audit experience has been negatively correlated with judgmental consensus. For example, Hamilton and Wright [1982] found the mean correlation for experienced auditors' internal control evaluations to be .71, while the mean correlation for inexperienced auditors was .73. Joyce [1976] also reported that, in his study, the level of consensus decreased as experience of practitioners increased. The reason for this decrease in consensus is unclear, although various possibilities exist. For instance, it is feasible that the limited audit planning experience of seniors results in their use of simplifying heuristics in audit planning judgments which results in higher consensus. A second possibility is that seniors may be less confident in audit planning, and thus less willing to vary their responses from case to case, than are managers. Furthermore, managers

would generally have a broader range of auditing experiences than seniors, which could further contribute to the lower level of consensus among managers.

The mean CNS_{MAD} for managers' internal control evaluations was 2.627 with a standard deviation of 1.242 and a range of 0.625 to 6.375. For seniors, the mean was 2.769 with a standard deviation of 1.027 and a range of 0.750 to 5.125. Comparing the mean CNS_{MAD} of seniors' internal control evaluations (2.769) with that of students (2.396) reveals that students agree more among themselves than do seniors concerning the strength of control structures. This difference in students' and seniors' CNS_{MAD} 's (-0.373) is statistically significant ($P = .008$). One possible explanation for this is that students may have utilized some type of simplifying heuristic such as equal weighting of cues when evaluating controls which neither seniors nor managers used. To test this possibility, separate ANOVAs for students', seniors', and managers' internal control evaluations were calculated. An examination of the eta-squared index for each of the four factors in the ANOVA based on students' responses led to the rejection of this possibility. Students' eta-squared indices ranged from .119 for the "Monthly Statements Reviewed?" factor, to .233 for the "Subsidiary Ledger Reconciled by Client?" factor. In comparison, eta-squared indices for seniors ranged from .047 for the "Monthly Statements Reviewed?" factor, to .211 for the "Billings Separate from Cash Receipts?" factor. For managers, the range was from .094 for the "Write-offs Approved?" factor to .183 for the "Subsidiary Ledger Reconciled by Client?" factor. Thus, there is no evidence that students employed an equal-weighting heuristic to any greater extent than seniors or managers. A second possibility is that,

since students had recently covered the topic of internal control evaluation in class, many of them may have used similar evaluation processes and heuristics, whereas practitioners, with their broader range of auditing experiences, may have employed judgment strategies which differed from one another to a greater extent than those of students. It is also interesting to note that the seniors had greater consensus among themselves when planning the extent of testing than when evaluating internal controls ($2.12 < 2.77$), while the students showed more consensus among themselves when evaluating controls than when planning the extent of testing ($2.86 > 2.40$).

A comparison of the mean CNS_{MAD} for internal control evaluations of seniors (2.665) with that of managers (2.627) suggests that seniors agree less among themselves concerning the strength of the control structure than do managers⁴. However, the magnitude of this difference (0.038) is not statistically significant ($P > .10$).

To further examine the effects of audit experience on audit judgment, two 2×2 ANOVAs were performed; one based on the judgments of students and seniors, and the other based on the judgments of seniors and managers. The dependent variable in each ANOVA was CNS_{MAD} ; the independent variables were experience level and audit task. The results of the ANOVAs are presented in Tables 4.19 and 4.20. As indicated in Table 4.19, which presents the ANOVA results based on the judgments of students and seniors, the experience-level $\times$ audit-task interaction was significant ($P < .001$). The interaction is plotted in Figure 4.11.

The observed results are consistent with the hypothesized effect. Auditing students appear to do reasonably well when evaluating the relative strength of various internal control structures. Apparently,

audit experience does not play a major role in this particular audit task. However, in the area of audit planning, students' judgments appeared to differ significantly from those of experienced audit seniors, indicating the importance of the role of audit experience in the development of the skills necessary for audit planning. While students may be acceptable surrogates for practitioners in internal control evaluation tasks, students appear to lack the skills necessary in less-structured tasks such as audit planning.

Table 4.19
Inter-Auditor Consensus (CNS_{MAD})
Students and Seniors
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	3.05	1	3.048	3.20	.075
S/A (Error)	169.70	178	0.953		
Audit Task (B)	0.74	1	0.745	0.77	
A x B	27.99	1	27.987	28.78	.000 **
B x S/A (Error)	173.07	178	0.972		

** P<.01

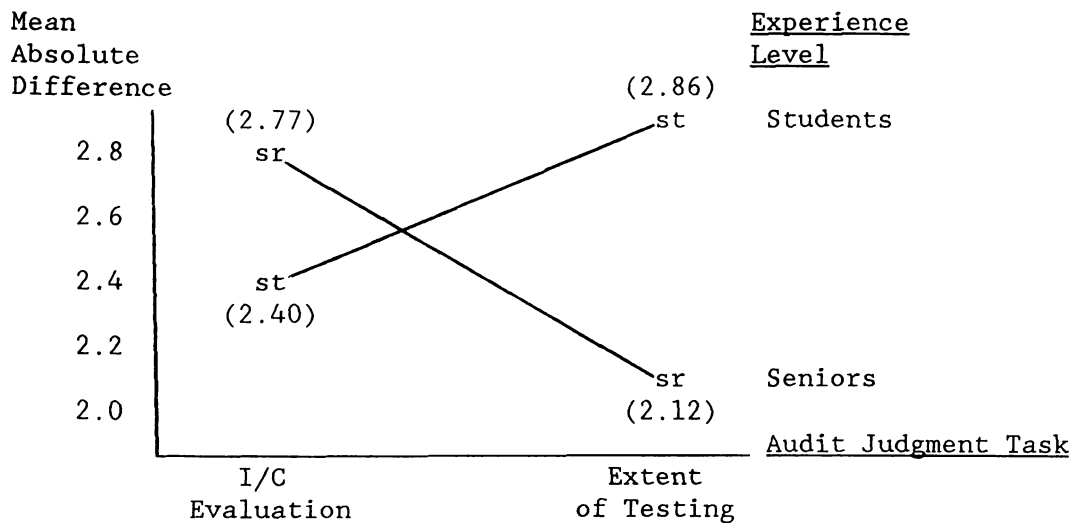


Figure 4.11
Experience Level x Audit Task Interaction (CNS_{MAD})
Students and Seniors

The difference between students' and seniors' CNS_{MAD} 's when planning the extent of testing ($2.86 - 2.12 = 0.74$), is larger than the difference in their MADs when evaluating internal control ($2.40 - 2.77 = -0.37$). This parallels the results observed in Table 4.13 based on judgments of students and all practitioners, and is consistent with hypothesis eight.

Table 4.20 presents the ANOVA based on the judgments of seniors and managers.

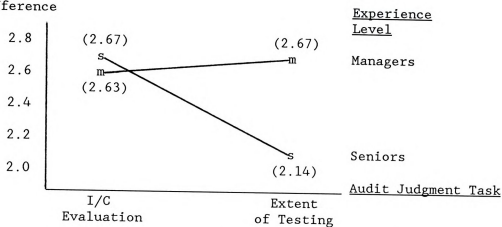
Table 4.20
Inter-Auditor Consensus (CNS_{MAD})
Seniors and Managers
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	3.41	1	3.407	3.43	.066
S/A (Error)	109.27	110	0.993		
Audit Task (B)	3.35	1	3.345	3.63	.059
A x B	4.54	1	4.536	4.93	.028 *
B x S/A (Error)	101.27	110	0.921		

* .01 < P < .05

The experience-level x audit-task interaction is significant ($P = .028$). This interaction is plotted in Figure 4.12. The pattern of the interaction is interesting. As expected, there was not a significant difference between managers' and seniors' consensus levels in the internal control evaluation task. However, in the audit planning task, *seniors* showed significantly more consensus among themselves than did managers. Apparently, when performing ill-structured audit tasks, audit experience increases judgmental agreement among auditors in the early years of their careers, but beyond some point, the variety of experiences acquired by auditors causes their judgment processes to diverge.

Mean
Absolute
Difference



NOTE: The above CNS_{MAD} values for seniors differ somewhat from those in Figure 4.11 due to the random elimination of several CNS_{MAD} values for seniors. This was done to equalize the number of observations for seniors and managers (see endnote #4).

Figure 4.12
Experience Level x Audit Task Interaction (CNS_{MAD})
Seniors and Managers

In summary, the current study has found that professional experience in the early years of auditors' careers tends to increase the level of judgmental agreement among auditors. In the early stages of auditors' careers, experience appears to educate the auditors in the manner in which their firm expects them to perform audit tasks. This education produces a convergence in judgment strategies. However, as auditors continue to gain more experience, the diversity of experiences acquired by different auditors appears to result in decreasing levels of judgmental agreement across auditors. This observation is consistent with the results of previous studies. Nanni [1984] describes the effects of experience in the following way:

Audit experience is assuredly associated with some learning effects, which might account for increased consensus with increasing experience levels. Relatively new auditors may also tend to be overly conservative in their evaluations

until they become confident in their judgment. Beyond some threshold, however, such experience differences should begin to diminish. At that point, experience may become secondary to responsibilities or evaluation goals in determining inter-auditor differences. ... The greater uncertainty they (managers) face from their broader perspectives may reduce consensus.

4.3.5 G, C, & R for Seniors and Manager

In order to better understand the sources of the differences in the judgment processes of students, seniors, and managers, G_{MAD} , R_{MAD} , and C_{MAD} were calculated separately for seniors and managers. The results, along with G_{MAD} , R_{MAD} , and C_{MAD} previously calculated for students, are presented in Table 4.21.

Table 4.21
Mean Values for G_{MAD} , R_{MAD} , and C_{MAD}
Managers, Seniors, & Students

	G_{MAD}	R_{MAD}	C_{MAD}
Managers -- Audit Planning	2.415 (0.871)	0.819 (0.422)	1.109 (0.482)
Seniors -- Audit Planning	1.957 (0.692)	0.722 (0.269)	0.953 (0.320)
Students -- Audit Planning	2.669 (1.361)	0.822 (0.269)	1.054 (0.381)
Managers -- I/C Evaluation	2.615 (1.315)	0.793 (0.287)	1.094 (0.405)
Seniors -- I/C Evaluation	2.597 (1.045)	0.676 (0.440)	1.020 (0.539)
Students -- I/C Evaluation	2.096 (0.880)	0.869 (0.526)	1.156 (0.500)

Standard deviations are shown in parentheses.

It appears from Table 4.21 that the major source of differences between the responses of auditors at different levels of audit experience arises from differences in cue weighting agreement. Differences in judgment

consistency, R_{MAD} , and differences in the degree of configurality in auditors' judgment processes, C_{MAD} , appear to have little influence. To test this, two sets of three 2 x 2 ANOVAs were performed. The first set consisted of three ANOVAs in which G_{MAD} , R_{MAD} , and C_{MAD} for students and seniors were the dependent variables, and experience level and audit task were the factors. The second set of ANOVAs was based on G_{MAD} , R_{MAD} , and C_{MAD} for seniors and managers.

The results of the ANOVAs based on judgments of students and seniors are presented in Tables 4.22, 4.23, and 4.24. As shown in Table 4.22, the experience-level x audit-task interaction was significant ($P < .001$), indicating that the extent to which students and seniors differ in their judgments of how different types of evidence should be weighted depends on the audit task being performed.

Table 4.22
Cue-Weighting Agreement (G_{MAD})
Students and Seniors
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	1.00	1	0.996	0.94	
S/A (Error)	188.65	178	1.060		
Audit Task (B)	0.10	1	0.102	0.10	
A x B	33.12	1	33.116	31.90	.000 **
B x S/A (Error)	184.80	178	1.038		

** $P < .01$

The nature of the experience-level x audit-task interaction is plotted in Figure 4.13. As illustrated in Figure 4.13, seniors have a higher degree of cue-weighting agreement among themselves when planning the extent of audit testing than do students, and that the opposite is true when evaluating internal controls.

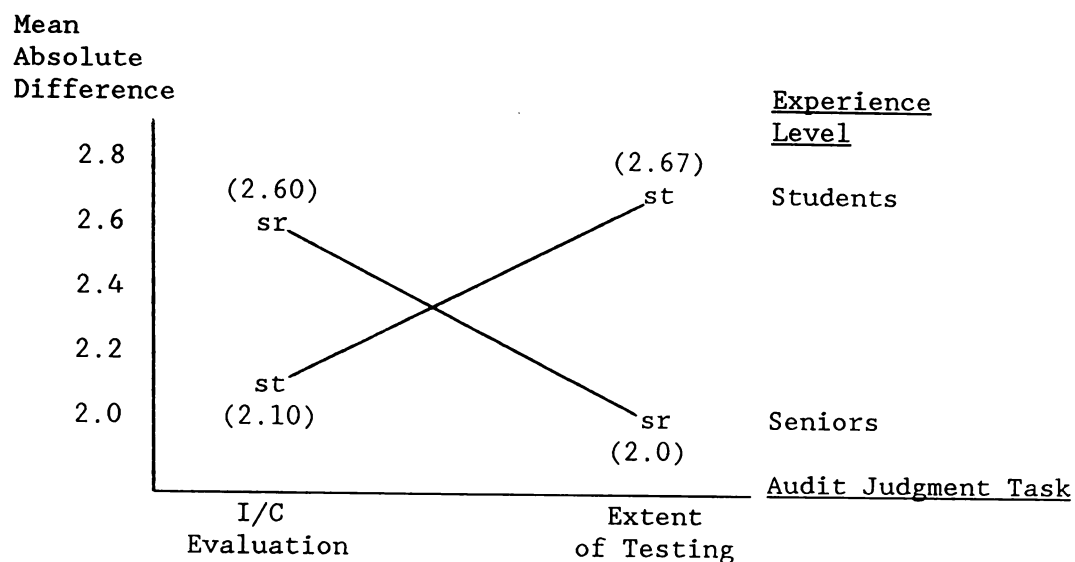


Figure 4.13
Experience Level x Audit Task Interaction (G_{MAD})
Students and Seniors

Table 4.23 presents the ANOVA for judgment consistency (R_{MAD}) of students and seniors.

Table 4.23
Judgment Consistency (R_{MAD})
Students and Seniors
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	0.43	1	0.430	2.37	.132
S/A (Error)	6.90	38	0.181		
Audit Task (B)	0.00	1	0.000	0.00	
A x B	0.04	1	0.043	0.35	
B x S/A (Error)	4.77	38	0.126		

There were no significant effects for R_{MAD} . The experience-level x audit-task interaction for CNS_{MAD} could not have been caused by the student-senior difference in R_{MAD} since the difference between seniors' R_{MAD} 's and students' R_{MAD} 's was nearly twice as large for internal control evaluation than for audit planning, although this was not large enough to produce an interaction. Thus, differences in the consistency

with which students and audit seniors make judgments are not responsible for the experience-level x audit-task interaction in consensus.

The results of the ANOVA for configural cue utilization agreement (C_{MAD}) are presented in Table 4.24.

Table 4.24
Configural Cue Utilization Agreement (C_{MAD})
Students and Seniors
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	1.26	1	1.262	5.67	.018 *
S/A (Error)	39.61	178	0.223		
Audit Task (B)	0.64	1	0.641	3.74	.054
A x B	0.03	1	0.028	0.16	
B x S/A (Error)	30.52	178	0.171		

At the conventional alpha level of .05, experience level was the only significant effect for C_{MAD} ($P = .018$) and resulted from seniors having more agreement on configurality than students in both tasks. The experience-level x audit-task interaction for CNS_{MAD} could not have been caused by the student-senior difference in C_{MAD} since the difference between seniors' C_{MAD} 's and students' C_{MAD} 's was larger in internal control evaluation than in audit planning, but not large enough to produce an interaction. Differences in the extent to which audit seniors and students employ configurality in their judgment processes is not responsible for the experience-level x audit-task interaction in consensus. Therefore, the significant experience-level x audit-task interaction in auditor consensus (CNS_{MAD}), in which the difference between students' and seniors' audit planning consensus (0.742) is greater than the difference between students' and seniors' internal control evaluation consensus (-0.373), is apparently due primarily to

differences in cue weighting agreement (G_{MAD}).

Tables 4.25, 4.26, and 4.27 present the results of the ANOVAs based on G_{MAD} , R_{MAD} , AND C_{MAD} for seniors and managers.

Table 4.25
Cue-Weighting Agreement (G_{MAD})
Seniors and Managers
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	4.24	1	4.239	3.93	.049 *
S/A (Error)	118.68	110	1.079		
Audit Task (B)	5.73	1	5.726	5.83	.017 *
A x B	1.34	1	1.338	1.36	.245
B x S/A (Error)	108.03	110	0.982		

* .01<P<.05

Table 4.26
Judgment Consistency (R_{MAD})
Seniors and Managers
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	0.12	1	0.118	0.91	
S/A (Error)	3.88	30	0.129		
Audit Task (B)	0.02	1	0.024	0.16	
A x B	0.01	1	0.006	0.04	
B x S/A (Error)	4.66	30	0.155		

Table 4.27
Configural Cue Utilization Agreement (C_{MAD})
Seniors and Managers
ANOVA Table

Source	SS	df	MS	F	P
Experience Level (A)	1.01	1	1.013	5.83	.017 *
S/A (Error)	19.10	110	0.174		
Audit Task (B)	0.00	1	0.004	0.02	
A x B	0.00	1	0.000	0.00	
B x S/A (Error)	23.99	110	0.218		

* .01<P<.05

Both experience level and audit task had significant main effects for G_{MAD} ($P = 0.049$ and $P = 0.017$, respectively). Seniors agreed more with one another on cue-weighting agreement than did managers in both audit tasks, and seniors and managers both showed greater cue weighting agreement when planning the extent of testing than when evaluating internal control. The senior-manager difference in audit planning cue weighting agreement (0.458) was larger than that in internal control evaluation (0.018), although not large enough to create an interaction ($P = .245$).

There were no significant effects for R_{MAD} . The senior-manager difference in audit planning judgment consistency (0.097) was smaller than that in internal control evaluation (.117). Therefore, the experience-level x audit-task interaction in CNS_{MAD} for seniors and managers could not have been caused by senior-manager differences in judgment consistency.

The only significant effect for C_{MAD} was experience level ($P = 0.017$). Seniors agreed more with one another concerning agreement on configurality than did managers in both tasks. The senior-manager difference in audit planning configurality agreement (0.156) was larger than that in internal control evaluation (0.074), but not large enough to create an interaction. Thus, the decomposition of CNS_{MAD} for seniors and managers into G_{MAD} , R_{MAD} , and C_{MAD} provides no clear indication of the source of the experience-level x audit-task interaction in which the manager-senior difference in audit planning CNS_{MAD} 's (0.539) was larger than that in internal control evaluation (-0.142). However, it is clear the manager-senior differences in judgment consistency are not responsible for the interaction. The observed results suggest that the

interaction may be due to a combination of differences in cue-weighting agreement and configurality agreement -- differences which, individually, were not statistically significant.

4.4 AVERAGING- VERSUS SUMMING-TYPE EVIDENCE INTEGRATION

One goal of this research is to investigate the manner in which auditors integrate the results of tests of individual internal controls when evaluating the effectiveness of the internal control structure. This phase of the research examines the nature of the evidence integration functions used by auditors when making this evaluation. The first step in identifying the nature of the integration function is to determine whether auditors tend to employ an additive process or whether their integration process tends to be nonadditive. Previous research has found very little evidence of configurality in auditors' judgment processes involved in the evaluation of controls. The conclusion has generally been that auditors employ an additive evidence integration process when evaluating controls.

Identifying the integration function as additive or nonadditive is an important initial step in determining the form of the functional integration rule used by auditors when evaluating the control structure. However, previous research has left some ambiguity in the form of the integration rule. To date, no study has been able to identify whether this evaluation process involves an averaging or a summing of evidence, both of which are additive processes. Since the test of additive versus nonadditive integration does not distinguish between averaging- versus summing-type integration, additional tests are needed to distinguish between them. One objective of the current study is to identify whether

auditors evaluate controls through an averaging process or through a summing process. The steps involved in identifying the nature of the integration process is illustrated in Figure 4.14.

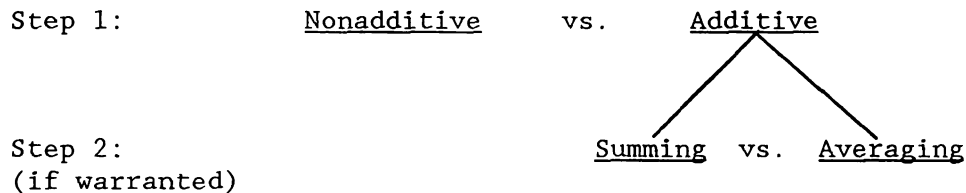


Figure 4.14
Steps in Identifying Summing vs. Averaging
Integration Processes

As mentioned previously, the first step is to distinguish between additive and nonadditive evidence integration. This was done by performing a 3×2^4 ANOVA for subjects' internal control evaluations. Factor A was the subjects' experience level (student, audit senior, or manager). Factor B dealt with whether the client reconciles the accounts receivable subsidiary ledger with the general ledger at least monthly (yes, no). Factor C was whether billings are performed by personnel who are independent of credit and cash receipts personnel (yes, no). Factor D dealt with whether customers' monthly statements are reviewed and mailed by personnel who are independent of those who maintain the accounts receivable subsidiary ledger (yes, no). Factor E was concerned with whether write-offs of receivables are approved by personnel who are independent of cash receipts and the credit manager (yes, no). Table 4.28 presents the cell means and standard deviations for subjects' control ratings, averaged across experience level, and cell means across experience level. Table 4.29 presents the ANOVA table for control ratings.

Table 4.28
Internal Control Evaluation Ratings
Cell Means and Standard Deviations

Panel A
Collapsed Across Experience Level

		Subsidiary Ledger Reconciled?				
		<u>No</u>	<u>Yes</u>			
		Monthly Statements Reviewed?		Monthly Statements Reviewed?		
		<u>No</u>	<u>Yes</u>	<u>No</u>	<u>Yes</u>	<u>Averages</u>
Billings Independent of Cash <u>Receipts</u>	Write-offs	8.63	11.33	12.08	14.54	11.65
	<u>Approved</u>	(2.52)	(2.78)	(2.13)	(1.62)	
						10.04
	Write-offs	5.92	8.42	8.63	10.75	8.43
Billings NOT Indep- endent of <u>Cash Receipts</u>	<u>NOT Approved</u>	(2.23)	(1.91)	(3.01)	(2.51)	
						8.20
	Write-offs	5.08	7.71	8.54	11.46	
	<u>Approved</u>	(2.08)	(2.85)	(2.23)	(1.93)	7.02
	Write-offs	3.38	4.92	6.71	8.33	5.84
	<u>NOT Approved</u>	(2.08)	(1.77)	(3.54)	(3.10)	
	<u>Averages</u>	5.75	8.10	8.99	11.27	
		6.93		10.13		8.53

Table 4.28 (cont'd.)

Panel B
Grouped by Experience Level

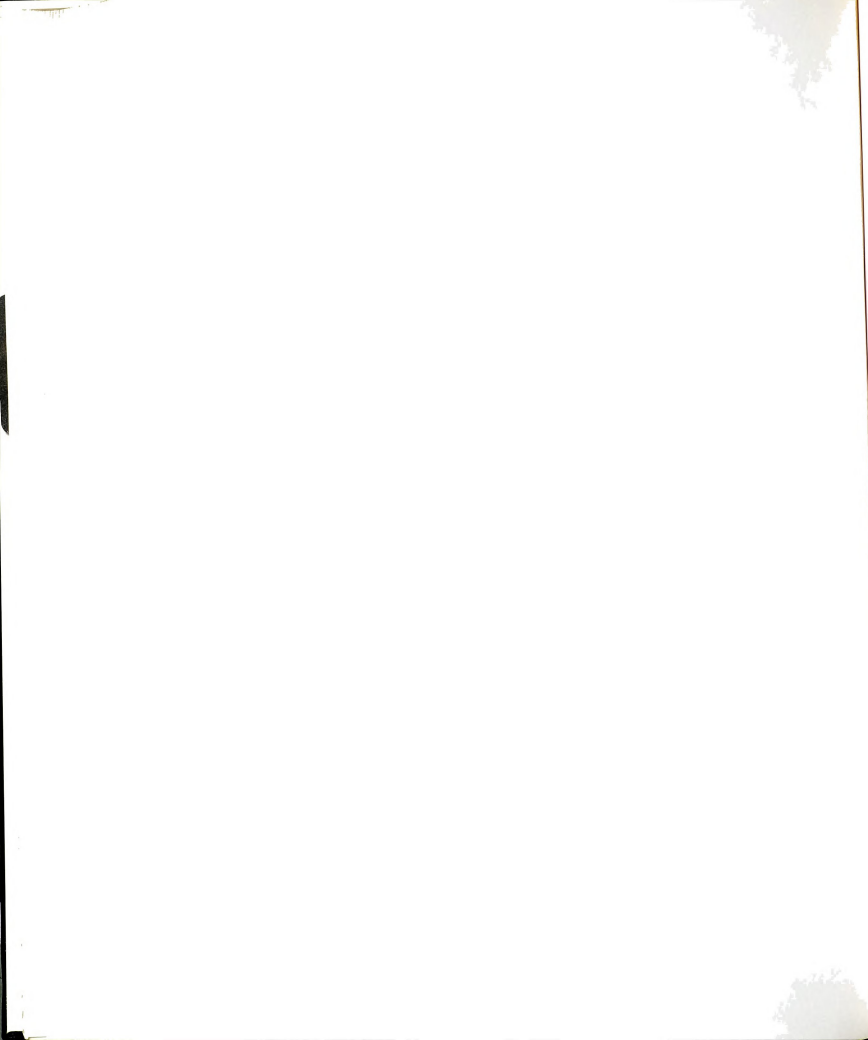
<u>Students</u>	<u>Seniors</u>	<u>Managers</u>
7.82	8.98	8.77
(3.78)	(3.74)	(3.65)

Table 4.29
Internal Control Evaluation Ratings
ANOVA Table

Source	SS	df	MS	F	P	Eta ²
Experience Level (A)	98.47	2	49.237	2.38	.104	.018
Subsid. Ldgr Recncl'd? (B)	988.17	1	988.167	47.83	.000 **	.183
A x B	8.85	2	4.424	0.21		.002
S/AB	867.75	42	20.661			
Billings Indpndt? (C)	876.04	1	876.042	147.73	.000 **	.163
A x C	6.79	2	3.393	0.57		.001
B x C	7.59	1	7.594	1.28	.264	.001
A x B x C	6.52	2	3.258	0.55		.001
C x S/AB	249.06	42	5.930			
Stmts Reviewed? (D)	513.38	1	513.375	175.57	.000 **	.095
A x D	27.39	2	13.695	4.68	.014 *	.005
B x D	0.09	1	0.094	0.03		.000
A x B x D	13.33	2	6.664	2.28	.114	.002
D x S/AB	122.81	42	2.924			
C x D	1.76	1	1.760	0.95		.000
A x C x D	5.35	2	2.674	1.44	.248	.001
B x C x D	1.50	1	1.500	0.81		.000
A x B x C x D	0.89	2	0.445	0.24		.000
C x D x S/AB	78.00	42	1.857			
Write-offs Approved? (E)	748.17	1	748.167	87.85	.000 **	.139
A x E	13.76	2	6.878	0.81		.003
B x E	6.51	1	6.510	0.76		.001
A x B x E	53.38	2	26.690	3.13	.053	.010
E x S/AB	357.69	42	8.516			
C x E	17.51	1	17.510	7.84	.007 **	.003
A x C x E	14.69	2	7.346	3.29	.047 *	.003
B x C x E	2.04	1	2.042	0.91		.000
A x B x C x E	1.01	2	0.503	0.23		.000
C x E x S/AB	93.75	42	2.232			
D x E	12.76	1	12.760	6.24	.016 *	.002
A x D x E	6.44	2	3.221	1.58	.218	.001
B x D x E	0.17	1	0.167	0.08		.000
A x B x D x E	3.76	2	1.878	0.92		.001
D x E x S/AB	85.88	42	2.045			
C x D x E	5.04	1	5.042	3.00	.090	.001
A x C x D x E	8.69	2	4.346	2.59	.087	.002
B x C x D x E	0.01	1	0.010	0.01		.000
A x B x C x D x E	10.19	2	5.096	3.03	.058	.002
C x D x E x S/AB	70.56	42	1.680			
Totals	5385.75	383				

* .01<P<.05

** P<.01



Two 2-way interactions involving elements of the control structure were present. The nature of these interactions is illustrated in Figures 4.15 & 4.16.

Evaluation of
Strength of
Internal Control
Structure

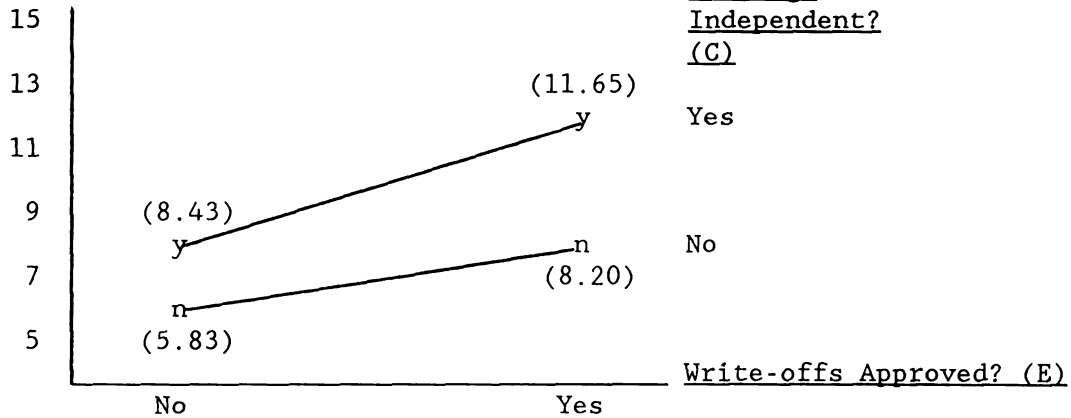


Figure 4.15
Internal Control Evaluation Ratings
C x E Interaction

Evaluation of
Strength of
Internal Control
Structure

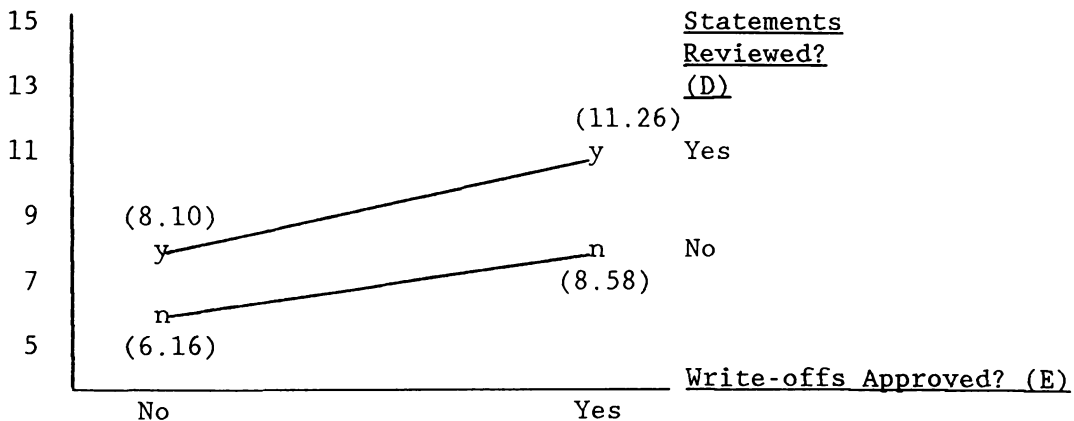


Figure 4.16
Internal Control Evaluation Ratings
D x E Interaction

As illustrated in Figures 4.15 and 4.16, when one of the two factors is negative, i.e., when one of the two controls is not operating, the effect of the other factor is attenuated. For example, when Factor E is at level "Yes", a change in Factor C from level "No" to level "Yes" results in an increase in the control rating of 3.45. However, when Factor E is "No", the same change in Factor C results in a smaller increase of only 2.60. These results indicate that auditors' judgment processes involved in the evaluation of the internal control structure tend to be *nonadditive*. Apparently, auditors do not consider the results of tests of individual internal controls independently. Rather, the effects of a particular individual control on auditors' evaluations of the control structure are influenced by the results of tests of other individual controls. This implies that auditors employ a more complicated evaluation process when evaluating controls than previously believed. Auditors' professional judgments apparently go beyond a summing or averaging of individual factors. Further research into the nature of auditors' evidence integration processes is deemed warranted. One suggestion would be to use statistical techniques which are better suited than the traditional ANOVA at discovering, when present, the particular types of interactive process auditors appear to be using.

4.5 ADDITIONAL ANALYSES

A debriefing questionnaire was included at the end of Booklet #3 to gather information about the subjects' educational and professional backgrounds and about their attitudes toward auditing and the experiment.

Eighty percent of the audit seniors and 72% of the managers found the cases to be either very realistic or somewhat realistic. Sixty percent of the seniors and 56% of the managers reported they would *not* have found additional information helpful in completing the cases. These results suggest there was a reasonably high degree of realism in the experiment's requirements and decision environment. Seventy percent of the seniors and 67% of the managers found the experiment to be either very interesting or fairly interesting. Since subjects were not explicitly rewarded for "good" performance when evaluating the experimental cases, the researcher believes it is important that the subjects find the task intrinsically interesting. By nature, people are likely to be more attentive and diligent when performing tasks they find interesting than when performing tasks they find boring. It appears that, in general, subjects found the task interesting.

On average, managers took 60 minutes to complete the experiment. Seniors reported an average time of 58 minutes, and students completed the experiment in an average of 36 minutes. When planning the experiment, one feature which the researcher deemed desirable was that the experimental task take no longer than two hours for the subjects to complete. It is believed that the validity of the results of an experiment may suffer as subjects become fatigued and less attentive to the task. In addition, it was believed that the experiment's response rate might suffer if the task were too lengthy. The experiment, on average, took less than one hour to complete.

Additional information about subjects' willingness to accept risk, extensiveness of audit procedures performed in relation to other auditors, and accuracy of time estimates was gathered in the debriefing



questionnaire. If results of the audit planning experiment had been insignificant, then this additional information would have been used to stratify subjects in an attempt to reduce the between-subjects response variability entering the ANOVA's error term. However, given the significance of the results of the research, no further analyses were performed.

ENDNOTES

1. In all analyses, cell sizes were equalized by randomly eliminating observations. The experiment had been planned so as to gather complete responses from 20 students, 20 seniors, and 20 managers. While all 20 students and all 20 seniors provided complete responses, only 19 managers provided complete responses to the audit planning experiment, and only 18 managers provided complete responses to the internal control evaluation experiment. Therefore, observations for one student and one senior, selected randomly, were eliminated.

Although power is reduced when observations are eliminated, the use of equal cell sizes was deemed desirable for three reasons. First, equal cell sizes results in each experimental condition receiving equal weight in the analysis. Second, equal cell sizes minimizes the effects of violations of the assumptions of normality and homogeneity of variance. Bartlett's test indicated the presence of significant heterogeneity of variance in subjects' responses. Therefore, equality of cell sizes was considered to be important in the analyses. Third, equal cell sizes simplify the calculation of sums of squares in the ANOVAs (Keppel [1982] pg. 99).

2. Errors terms for each interaction contrast were based on only the data involved in the contrast, as opposed to using the error term from the overall ANOVA. This was deemed desirable since Bartlett's test indicated heterogeneity of error variance across treatment conditions.

3. A numerical example, using four cases, will illustrate the way in which the observed results are consistent with the audit risk model:

$$\text{Audit Risk Model: } \text{DDR} = \frac{\text{AAR}}{\text{IR} \times \text{CR} \times \text{AnR}}$$

where: DDR = Desired Detection Risk
 AAR = Acceptable Audit Risk
 IR = Inherent Risk
 CR = Control Risk
 AnR = Analytical Risk

In case 1 below, assume the internal operating environment is unfavorable; i.e., inherent risk = 1.0. In case 2, assume the internal operating environment is more favorable; i.e., inherent risk = .60. Assume acceptable audit risk and analytical risk in both cases is .05 and 1.0, respectively.

Case 1:

High Inherent Risk; Low Control Risk; High Analytical Risk:

$$\text{DDR} = \frac{.05}{1.0 \times .3 \times 1.0} = .166.$$

High Inherent Risk; High Control Risk; High Analytical Risk:

$$\text{DDR} = \frac{.05}{1.0 \times .9 \times 1.0} = .055.$$

In case 1, a decrease in the strength of controls, i.e., a change in control risk from .3 to .9, results in a decrease in desired detection risk (which results in an increase in audit testing) of .111 (.166 - .055):

Case 2:

Low Inherent Risk; Low Control Risk; High Analytical Risk:

$$\text{DDR} = \frac{.05}{.6 \times .3 \times 1.0} = .277.$$

Low Inherent Risk; High Control Risk; High Analytical Risk:

$$\text{DDR} = \frac{.05}{.6 \times .9 \times 1.0} = .093.$$

In case 2, the same decrease in the strength of controls results in a larger decrease in desired detection risk (and a larger increase in audit testing) of .184 (.277 - .093):

The difference in the magnitude of the change in DDR in the Low Inherent Risk case (Case 2) and the change in DDR in the High Inherent Risk case (Case 1) is .073 (.184 - .111).

Case 3:

High Inherent Risk; Low Control Risk; Low Analytical Risk:

$$\text{DDR} = \frac{.05}{1.0 \times .3 \times .4} = .417.$$

High Inherent Risk; High Control Risk; Low Analytical Risk:

$$\text{DDR} = \frac{.05}{1.0 \times .9 \times .4} = .139.$$

In case 3, a decrease in the strength of controls, i.e., a change in control risk from .3 to .9, results in a decrease in desired detection risk (which results in an increase in audit testing) of .278 (.417 - .139):

Case 4:

Low Inherent Risk; Low Control Risk; Low Analytical Risk:

$$\text{DDR} = \frac{.05}{.6 \times .3 \times .4} = .694.$$

Low Inherent Risk; High Control Risk; Low Analytical Risk:

$$\text{DDR} = \frac{.05}{.6 \times .9 \times .4} = .231.$$

In case 4, the same decrease in the strength of controls results in a very large decrease in desired detection risk (and a very large increase in audit testing) of .463 (.694 - .231):

The difference in the magnitude of the change in DDR in the Low Inherent Risk; Low Analytical Risk case (Case 4) and the change in DDR in the High Inherent Risk; Low Analytical Risk case (Case 3) is .185 (.463 - .278). This is more than twice the corresponding difference in the High Analytical Risk cases (Cases 1 and 2). This same pattern of results is observed in the C x D and C x D x E interactions. Thus, the nature of the observed C x D and C x D x E interactions is consistent with the audit risk model.

4. In the comparison of students' and seniors' consensus levels, responses for all 20 students and all 20 seniors were used. However, in the comparison of seniors' and managers' consensus levels, several pairwise observations for seniors' CNS_{MAD} were randomly eliminated to equalize cell sizes for managers and seniors. This resulted in a change in CNS_{MAD} from 2.121 to 2.136 for seniors' audit planning judgments and in a change from 2.769 to 2.665 for seniors' internal control evaluations.

CHAPTER 5 - SUMMARY, CONCLUSIONS, AND SUGGESTIONS FOR FUTURE RESEARCH

5.1 SUMMARY OF EXPERIMENTAL FINDINGS

5.1.1 Budgeted Audit Hours Judgments

One goal of this research was to investigate the manner in which various factors influence auditors' judgments of the appropriate extent of tests of details. An experiment was performed to investigate this issue. Two methods were used in the experiment to measure subjects' planned extent-of-testing judgments. One method involved asking subjects to indicate how many hours they would budget for tests of details of the accounts receivable balance. The second method involved asking subjects to indicate graphically their planned extent of tests of details of the accounts receivable balance by placing a slash mark on a four-inch line anchored "no testing of receivables" on the left and "extensive testing of receivables" on the right. The purpose of this second extent-of-testing measure was to provide a dependent measure with less between-subject response variability than was likely to be present in the "budgeted audit hours" variable. Excessive between-subject response variability in the reported "budgeted audit hours" could mask the presence of significant main effects or interactions, when analyzing the subjects' responses. The results of the analyses of subjects' budgeted audit hours responses are summarized in this section. A summary of the analyses of subjects' extent-of-testing ratings is presented in Section 5.1.2.

Hypothesis one predicted a *direct* relationship between auditors' planned extent of tests of details of balances and conditions in the

clients' internal operating environment which increase the likelihood of the occurrence of material misstatements. The main effect for the client's internal operating environment was significant, supporting hypothesis one. This indicates that, when planning the audit, auditors' budgeted-audit-hours judgments are influenced systematically by conditions in the internal operating environment, with more testing deemed appropriate when conditions in the internal operating environment suggest an increased likelihood of the occurrence of material misstatements in the accounting records. Thus, the conclusion is that auditors, when planning the audit, do indeed take into consideration factors that affect the likelihood of the occurrence of errors or irregularities in the accounting records. This conclusion differs from that of Kaplan and Reckers' [1984] study which found *no* evidence that auditors' judgments are influenced by the integrity of management - a factor which the auditing literature states should affect the auditor's assessment of the likelihood that errors or irregularities may have occurred.

Hypothesis two predicted an *inverse* relationship between the planned extent of tests of details and the effectiveness of the internal control structure. The main effect for the internal control structure was significant, supporting hypothesis two. This indicates that auditors' budgeted-audit-hours judgments are influenced systematically by changes in the effectiveness of the internal control structure, with more hours budgeted when conditions in the internal control structure suggest an increased likelihood that internal control may not have prevented or detected material misstatements in the accounting records. Thus, the conclusion is that auditors, when planning the audit, do

indeed take into consideration the effectiveness of the internal control structure.

Hypothesis three predicted an *inverse* relationship between the planned extent of tests of details and the favorableness of the results of analytical procedures. The main effect for analytical procedures was significant, supporting hypothesis three. Therefore, it is concluded that auditors, when planning the appropriate extent of tests of details, do indeed take into consideration the favorableness of the results of analytical procedures, with more tests of details of balances deemed appropriate when analytical procedures suggest an increased likelihood of the presence of material misstatements in the accounting records, than when analytical procedures suggest a lower likelihood of misstatements.

Hypothesis four predicted a *direct* relationship between the planned extent of tests of details and the degree of volatility in the client's external business environment. The main effect for the client's external business environment was *not* significant. Accordingly, judgment must be reserved regarding whether conditions in the external environment influence auditors' extent-of-testing judgments. This issue is addressed further in Section 5.1.2.

Hypothesis five predicted that the effect of the internal control structure on auditors' planned extent of tests of details would depend on conditions in the client's internal operating environment. However, the internal-control x internal-operating-environment interaction was *not* significant when based on budgeted audit hours responses. These results do *not* indicate that the effect of the internal control structure on auditors' judgments concerning the appropriate number of

hours to budget for tests of details depends on conditions in the client's internal operating environment. Therefore, judgment must be reserved regarding the veracity of hypothesis five.

Hypothesis six predicted that the effect of the results of analytical procedures on auditors' planned extent of tests of details would depend on the effectiveness of the internal control structure. However, the analytical-procedures x internal-control interaction was *not* significant when based on budgeted audit hours responses. The results of this study do *not* indicate that the effect of analytical procedures on auditors' judgments concerning the appropriate number of hours to budget for tests of details depends on the effectiveness of the client's internal control structure. Therefore, judgment must be reserved regarding the veracity of hypothesis six.

Hypothesis seven predicted that the effect of the results of analytical procedures on auditors' planned extent of tests of details would depend on the degree of volatility in the client's external business environment. However, the analytical-procedures x external-environment interaction was *not* significant when based on budgeted audit hours responses. The results of this study do *not* indicate that the effect of analytical procedures on auditors' judgments concerning the appropriate number of hours to budget for tests of details depends on the degree of volatility in the client's external environment. Therefore, judgment must be reserved regarding the veracity of hypothesis seven.

The results of tests of hypotheses one through seven based on subjects' budgeted audit hours responses are summarized in Figure 5.1.

	<u>Supported</u>	<u>Not Supported</u>
Hypothesis One	X	
Hypothesis Two	X	
Hypothesis Three	X	
Hypothesis Four		X
Hypothesis Five		X
Hypothesis Six		X
Hypothesis Seven		X

Figure 5.1
Summary of Tests of Hypotheses One Through Seven
Based on Budgeted Audit Hours Responses

In the next section, hypotheses one through seven are reevaluated using subjects' extent-of-testing ratings rather than budgeted audit hours.

This was done since, in the preceding tests of four of the seven hypotheses, the lack of significance may have been due to low statistical power arising from high inter-subject variability in the budgeted audit hours responses. Extent-of-testing ratings had less between-subject response variability than budgeted audit hours responses.

5.1.2 Extent-of-Testing Ratings

Hypotheses one through seven were reevaluated using subjects' extent-of-testing ratings, rather than budgeted audit hours, as the dependent variable in the analyses. Consistent with the results observed for budgeted audit hours responses, the main effects for the client's internal operating environment, the internal control structure, and the results of analytical procedures were significant. These results are consistent with hypotheses one, two, and three. Apparently, auditors' extent-of-testing judgments are influenced by the internal operating environment, by the internal control structure, and by the results of analytical procedures. The extent of tests of details

increases when conditions in the internal operating environment, the control structure, or the results of analytical procedures suggest an increased likelihood of the presence of material misstatements in the accounting records.

The test of hypothesis four, when based on budgeted audit hours, revealed an *insignificant* main effect for the external environment. Thus, hypothesis four was not supported. However, when reevaluated based on extent-of-testing ratings, the main effect for the external environment was significant. Apparently, the greater inter-subject response variability in subjects' budgeted audit hours limited the power of the significance tests of the main effect for the external environment. This emphasizes the importance of careful selection of the dependent variable in behavioral research. The significant main effect for the external environment supports hypothesis four, and indicates that auditors' judgments of the appropriate extent of tests of details are influenced by the degree of volatility in the client's external business environment, with more tests of details of balances deemed appropriate when conditions in the environment indicate an increase in the likelihood that the client may experience financial trouble in the near future. When the client's external business environment is volatile, the auditor's business risk tends to increase. This, in turn, may increase the auditor's acceptable audit risk, and thus, the extent of audit testing. While these findings differ from those of Kaplan's [1985] study which did *not* conclude that there is a significant relationship between the extent of testing and environmental volatility, they are consistent with guidelines in the auditing literature which suggest that the client's business environment should be considered by

the auditor when planning the examination. However, given the level of significance ($P = .08$) of the main effect for environmental volatility, care should be exercised when drawing conclusions concerning the impact of environmental volatility on audit planning.

Hypothesis five predicted that a change in the effectiveness of the internal control structure would have a larger effect of auditors' extent-of-testing judgments when the internal operating environment is unfavorable, than when it is favorable. This hypothesis was based, in part, on the premise that effective controls would compensate somewhat for an internal environment which is conducive to a high likelihood of errors, and that an internal environment which is conducive to a low likelihood of errors would compensate for weak internal controls. A significant internal-operating-environment x internal-control interaction was observed. However, it was not of the expected form. Specifically, it was found that a change in the effectiveness of the control structure has its greatest effect on auditors' judgments when the internal operating environment is *favorable*, rather than unfavorable. This finding runs counter to form of the likelihood-of-errors x internal-control-strength interaction observed by Libby, Artman, and Willingham (LAW) [1985]. The interaction observed by LAW indicated that a change in control strength has its greatest effect on auditors' judgments when there is a *high* likelihood of errors occurring, rather than a low likelihood. Subjects in the LAW study were asked to indicate their control-reliance judgments, whereas the dependent variable in the current study was extent-of-testing ratings. It would seem to follow that the same set of conditions which produce a large change in auditors' control-reliance judgments should also produce a

large change in the extent of tests of details. For instance, a large decrease in control reliance should produce a larger increase in the extent of tests of details than would a small decrease in control reliance. However, the results of the current study indicate that LAW's results *cannot* be extended to auditors' extent-of-testing judgments. In the context of extent-of-testing judgments, a change in control strength has its greatest effect when the likelihood of errors is *low*, rather than high. The observed C x D interaction is thus opposite to the particular C x D interaction predicted by hypothesis five, and to the interaction observed by LAW. However, the observed interaction is consistent with the audit risk model.

Of additional interest is the fact that the internal-operating-environment x internal-control interaction occurs only when the results of analytical procedures are favorable; i.e., there is *no* internal-operating-environment x internal-control interaction when analytical procedures suggest a high likelihood of misstatements. The pattern of this three-way interaction between the internal operating environment, internal controls and analytical procedures indicates that when any of the three types of evidence suggest a *high* likelihood of misstatements, then the effects of the other evidence on auditors' extent-of-testing-of-details judgments are attenuated. Only when all three types of evidence are favorable is the auditor willing to reduce the extent of testing to a great degree. This is analogous to Ashton's [1974] observation that when auditors discover a weakness in internal controls, the results of other tests of the control structure seem to have little effect on the overall evaluation of the structure. Furthermore, these results are consistent with evidence in the psychology literature which

has found that negative information tends to be weighted heavier than positive information (Anderson & Alexander [1971], Hamilton & Zanna [1972], and Hodges [1974]). Finally, these observed results are consistent with the audit risk model (see endnote 3 in Chapter 4). The findings of this research thus lend empirical support to the audit risk model as a multiplicative function.

Hypothesis six predicted that the effect of analytical procedures on auditors' extent-of-testing-of-details judgments would depend on the effectiveness of the internal control structure. This is a prediction of an analytical procedures x internal control interaction. Unlike the results observed when subjects' budgeted audit hours responses were analyzed, this interaction is significant when based on extent-of-testing ratings. However, this interaction is significant *only* when conditions in the internal operating environment are favorable. When the internal operating environment is unfavorable, there is *no* interaction. Apparently, when conditions in the internal operating environment are such that there is a relatively high likelihood that material misstatements have occurred during the period under audit, the amount of audit assurance that auditors are willing to place on the results of analytical procedures is *not* influenced by the effectiveness of the internal control structure. Thus, the veracity of hypothesis six is conditional upon the state of the internal operating environment. When the internal environment is favorable, the results are consistent with hypothesis six. However, when the internal environment is unfavorable, hypothesis six is not supported. This conclusion is that when the internal operating environment, the internal control structure, or the results of analytical procedures suggest a *high* likelihood of

misstatements, then the effects of the other sources of information on auditors' extent-of-tests-of-details judgments are attenuated. This result, along with the pattern of the internal-operating-environment x internal-control interaction observed in the evaluation of hypothesis five and an unexpected interaction between the internal operating environment and analytical-procedures, suggests that the hypothesized two-way interactions do not adequately describe auditors' extent-of-testing judgment processes. Apparently three, rather than just two, factors interact simultaneously in their effects on auditors' extent-of-testing judgments. Therefore, when describing the interactive effects of any two of the three factors, the level of the third factor must also be considered. This indicates a highly configural judgment process is present in auditors' extent-of-testing judgments.

Hypothesis seven predicted that the effects of analytical procedures on auditors' extent-of-tests-of-details judgments would depend on the degree of volatility in the client's external business environment. This hypothesis was not supported. This research has found no evidence to support the statement in SAS 56 that the extent to which auditors rely on the results of analytical procedures depends on the stability of the external business environment. Future research should address this issue further.

The results of these tests of hypotheses one through seven based on subjects' extent-of-testing ratings are summarized in Figure 5.2.

5.1.3 Weighting of Evidence

In order to determine the relative importance, or weight, of each factor in subjects' judgments, the eta-squared index (η^2) was calculated for each factor. The internal operating environment was weighted more

	<u>Supported</u>	<u>Not Supported</u>
Hypothesis One	X	
Hypothesis Two	X	
Hypothesis Three	X	
Hypothesis Four	X	
Hypothesis Five	Involved in 3-way Interaction	
Hypothesis Six	Involved in 3-way Interaction	
Hypothesis Seven		X

Figure 5.2
Summary of Tests of Hypotheses One Through Seven
Based on Extent-of-Testing Ratings

heavily than any other factor in auditors' judgments of the appropriate extent of tests of details. Internal control and analytical procedures each received roughly half as much weight as the internal operating environment, but were nearly equal to one another in importance and were both significant factors in auditors' judgments. The external operating environment was less important, accounting for just over one percent of the variance in subjects' extent-of-testing ratings.

To examine whether subjects with different levels of audit experience weighed factors differently, eta-squared indices were calculated separately for students, seniors, and managers. Managers were influenced the most by the internal operating environment and placed much greater weight on it than did the seniors or students. While the internal operating environment was also the most important factor in seniors' judgments, seniors gave the internal environment just over half the weight that managers placed on it. Students gave even less weight to the internal operating environment than did seniors. Thus, the perceived relative importance of the internal operating environment appears to increase with audit experience.

The relative importance of analytical procedures appears to decline as audit experience increases. The results of analytical procedures is the most important factor in students' extent-of-tests-of-details judgments, but is a much less important factors in seniors' and managers' judgments. While SAS 58 has recently reemphasized the importance of analytical procedures, managers in particular apparently do not feel that analytical procedures are of much importance in the determination of the appropriate extent of tests of details.

The relative importance of the client's external business environment also appears to decline as audit experience increases. Students placed more weight on the external environment than did seniors or managers, and seniors placed slightly more weight on the external environment than did managers. In addition, all three subject groups - students, seniors, and managers - placed less weight on the external environment than any of the other three factors. Apparently, experienced auditors, as well as students, believe that the external business environment is of less importance in audit planning than the internal environment, the internal control structure, or the results of analytical procedures.

5.1.4 Effects of Professional Experience

The effects of professional experience on inter-auditor consensus were examined. Hypothesis eight predicted there would be a larger difference between the consensus levels of experienced and inexperienced auditors when planning the extent of testing than when evaluating the effectiveness of the internal control structure. This is a prediction of an experience-level x difficulty-of-audit-task interaction. Two methods were used to measure consensus: the Pearson product-moment

correlation coefficient and the mean absolute difference for pairwise judgments. When correlation was used as the measure of inter-auditor consensus, the interaction was not significant. Thus, hypothesis eight was not supported. Given the possibility that the non-significance of the interaction may have been due to a limitation of correlational measures in identifying differences in the judgments of pairs of auditors, hypothesis eight was reevaluated using mean absolute differences (CNS_{MAD}). Based on this second measure of consensus, the experience-level x difficulty-of-audit-task interaction was found to be significant. Practitioners showed more consensus than did students in the audit planning task, and there was a greater difference between practitioners' and students' consensus levels in the planning task than in the internal control evaluation task.

The objective of this phase of the research was to examine the effects of knowledge obtained through experience on audit planning judgments. The observed results are consistent with the hypothesized effect. Audit experience appears to have little effect on internal control evaluations. Apparently, auditing students develop sufficient skills through their course work to enable them to evaluate the relative strength of various internal control structures reasonably well. However, in the area of audit planning, audit experience appears to play a more vital role. Although students learn in their course work how various factors such as internal control effectiveness and the results of analytical procedures might affect the extent of tests of details, this task is less structured than internal control evaluation. While students may be acceptable surrogates for practitioners in internal control evaluation tasks, the results of this research suggest that

students are not good surrogates for practitioners in less-structured tasks such as audit planning.

5.1.5 Decomposition of Consensus

Three components of consensus were examined to determine which component is most affected by professional experience. The three components are cue weighting agreement, judgment consistency, and agreement on configural cue utilization. Results indicated that the previously observed significant experience-level x difficulty-of-audit-task interaction, in which the difference between students' and practitioners' audit planning consensus was greater than the difference between students' and practitioners' internal control evaluation consensus, was due primarily to differences in cue weighting agreement. Differences in the consistency with which auditors at different experience levels form their judgments were unable to account for the observed experience effects. Differences in inter-auditor agreement on configural cue utilization were likewise unable to account for the observed effects. The conclusion is that audit experience affects inter-auditor consensus mainly through its effect on cue weighting agreement.

5.1.6 Decomposition of Experience Levels

To obtain further insight into the effects of experience and how these effects differ across experience levels, the "practitioners" level of experience was partitioned into two groups -- seniors and managers. Comparisons were then made of the levels of judgmental agreement among students with those among seniors, and the levels of judgmental agreement of seniors with those of managers. When comparing the consensus levels of students with those of seniors, results were similar

to those observed when comparing students with all practitioners. That is, the experience-level x audit-task interaction was significant, with seniors showing significantly greater consensus than students in the audit planning task. Auditing students appear to do reasonably well when evaluating the relative strength of various internal control structures. Apparently, audit experience does not play a major role in this particular audit task. However, in the area of audit planning, students' judgments differed significantly from those of experienced audit seniors, indicating the importance of the role of audit experience in the development of the skills necessary for audit planning. Again, the conclusion is that while students may be acceptable surrogates for practitioners in internal control evaluation tasks, students appear to lack the necessary skills in performing ill-structured tasks such as audit planning.

When comparing seniors' consensus levels with those of managers, the experience-level x audit-task interaction was again significant. However, the pattern of the interaction suggested a different type of experience effect that had been previously observed. Specifically, in the audit planning task, *seniors* showed significantly more consensus among themselves than did managers. Apparently, when performing ill-structured audit tasks, audit experience increases judgmental agreement among auditors in the early years of their careers, but beyond some point, the variety of experiences acquired by auditors causes their judgment processes to diverge.

This study has found that professional experience in the early years of auditors' careers tends to increase the inter-auditor consensus. In the early stages of auditors' careers, experience appears

to educate the auditors in the manner in which they are to perform audit tasks. This education produces a convergence in judgment strategies. However, as auditors continue to gain more experience, the diversity of experiences acquired by different auditors appears to produce a divergence in judgment strategies and decreasing levels of inter-auditor consensus.

5.1.7 Averaging- Versus Summing-Type Evidence Integration

One goal of this research was to investigate the manner in which auditors integrate the results of tests of individual internal controls when evaluating the effectiveness of the internal control structure. The first step in this process was to identify whether auditors' integration processes tend to be additive or nonadditive. The results of previous research has found very little evidence of configurality, or nonadditivity, in auditors' evaluations of internal controls. Thus, the conclusion has generally been that auditors use an additive process.

Identifying the integration function as additive or nonadditive is an important initial step in determining the form of the functional integration rule used by auditors when evaluating the control structure. However, previous research has left some ambiguity in the form of the integration rule. To date, no study has examined whether this evaluation process involves an averaging or a summing of evidence, both of which are additive processes. The objective of this phase of the research was to examine this issue.

Results of the current study indicate that auditors' judgment processes involved in the evaluation of the client's internal control structure tend to be *nonadditive*. This was suggested by two 2-way interactions involving elements of the control structure. Since

evidence of a nonadditive function was found, further testing to distinguish between summing versus averaging processes was not warranted. Apparently, auditors do not consider the results of tests of individual internal controls independently. Rather, the effects of a particular individual control on auditors' evaluations of the control structure are influenced by the results of tests of other individual controls. This implies that auditors employ a more complicated judgment process when evaluating controls than previously believed, apparently going beyond a simple summing or averaging of individual factors.

5.2 LIMITATIONS OF THE STUDY

The experiment performed in this dissertation contains several of the same limitations which are present in much of the behavioral auditing research. These limitations tend to restrict the ability to generalize the results of this study to other groups of auditors and decision environments. First, strictly speaking, since subjects were not randomly selected from the entire population of auditors, the results cannot be generalized to other groups of auditors. Second, while this study attempted to provide subjects with realistic case situations, the realism in the auditors' decision environment was still limited. For instance, the fact that subjects were not allowed to discuss the cases with other auditors, the format of subjects' responses, and the format of the information which was provided to the subjects all caused the experimental decision environment to differ somewhat from auditors' actual decision environments. The fact that the experimental design did not consider the possibility of correlations between the experimental factors (e.g., in many audits, there may be a

positive correlation between the favorableness of analytical procedures and the effectiveness of the internal control structure) also may have caused the experimental environment to differ from auditors' actual decision environments. Finally, it should be noted that any single study is limited in its ability to allow us to draw conclusions regarding individuals' judgment processes. Replications of results are needed to assess the pervasiveness and robustness of results across different subjects and settings. These limitations should be kept in mind when interpreting the results of this study.

5.3 SUGGESTIONS FOR FUTURE RESEARCH

In the current research, only four factors which the auditing literature suggests may be relevant to audit planning judgments were examined. There are many other potentially relevant factors whose effects, if any, on auditors' judgments have not been investigated. Future research on audit planning might examine some of these factors and might consider alternate approaches to operationalizing relevant variables. For instance, in the audit planning task in the current study, the effectiveness of the internal control structure was manipulated by varying the effectiveness of several individual controls. Future research might consider manipulations of components of the other two elements of the control structure -- the control environment and the accounting process. Since the four variables of interest in this study were treated as aggregate variables, the effects of individual components of the variables (e.g., employee competence) on auditors' judgments cannot be evaluated. Future research might disaggregate the analytical procedures variable, and the two environmental variables to

examine the relative importance of the variables' components.

The audit risk model predicts that changes in analytical risk should combine multiplicatively with inherent risk and control risk in its influence on auditors' extent-of-testing judgments. The current study did not manipulate analytical risk, choosing instead to vary the actual results of analytical procedures. Future research should attempt to examine the effects of analytical risk on auditors' extent-of-detailed-testing judgments.

Future experimental research should also attempt to improve the external validity of experiments. To be able to generalize to larger groups of auditors other than those participating in the experiment, researchers should attempt to employ random selection processes when obtaining subjects. Furthermore, attempts should be made to increase the realism of the experiment's decision environment. For example, many audit decisions are made by audit teams, rather than by individual auditors working in isolation. By allowing groups of auditors to interact, researchers could not only improve the realism of the decision environment, but could also begin to study the effects of group discussion in audit planning.

Future research on the nature of auditors' evidence integration processes should be conducted. This study has found evidence of nonadditivity in auditors' judgments in two separate tasks -- extent-of-testing judgment and internal control evaluations. Future research should examine the pervasiveness of these results across other judgment tasks. Furthermore, this study draws no conclusions regarding the appropriateness or desirability of employing a nonadditive integration process, rather than an additive process, when determining the

appropriate extent of testing or when evaluating the internal control structure. Future research might examine whether the use of a particular integration process is more conducive to meeting the CPA firm's goals than other integration processes in various audit judgment settings.

LIST OF REFERENCES

LIST OF REFERENCES

- Abdolmohammadi, M., and A. Wright. "An Examination of the Effects of Experience and Task Complexity on Audit Judgments." The Accounting Review (January 1987): 1-13.
- Aly, H., and J. Duboff. "Statistical vs. Judgment Sampling: An Empirical Study of Auditing the Accounts Receivable of a Small Retail Store." The Accounting Review (January 1971): 119-128.
- American Institute of Certified Public Accountants. AICPA Professional Standards. Chicago: Commerce Clearing House, 1986.
- American Institute of Certified Public Accountants. Statement on Auditing Standards No. 55. New York, N.Y., 1988.
- American Institute of Certified Public Accountants. Statement on Auditing Standards No. 56. New York, N.Y., 1988.
- Anderson, N. "Averaging Versus Adding as a Stimulus-Combination Rule in Impression Formation." Journal of Experimental Psychology (1965), 70: 394-400.
- Anderson, N. "Averaging Model Analysis of Set-Size Effect in Impression Formation." Journal of Experimental Psychology (1967), 75: 158-165.
- Anderson, N. "Information Integration Theory Applied to Attitudes About U.S. Presidents." Journal of Educational Psychology (1973), 64: 1-8.
- Anderson, N. Foundations of Information Integration Theory. Academic Press, 1981.
- Anderson, N. Methods of Information Integration Theory. Academic Press, 1982.
- Anderson, N., and G. Alexander. "Choice Test of the Averaging Hypothesis for Information Integration." Cognitive Psychology (1971), 2: 313-324.
- Anderson, T., and M. Birnbaum. "Test of an Additive Model of Social Inference." Journal of Personality and Social Psychology (1976), 33: 655-662.

- Anderson, N., and C. Graesser. "An Information Integration Analysis of Attitude Change in Group Discussion." Journal of Personality and Social Psychology (1976), 34: 210-222.
- Anderson, N., and A. Jacobson. "Effect of Stimulus Inconsistency and Discounting Instructions in Personality Impression Formation." Journal of Personality and Social Psychology (1965), 2: 531-539.
- Anderson, N., R. Lindner, and L. Lopes. "Integration Theory Applied to Judgments of Group Attractiveness." Journal of Personality and Social Psychology, (1973), 26: 400-408.
- Arens, A., and J. Loebbecke. Auditing: An Integrated Approach. Fourth Edition, Prentice Hall, Inc., 1988
- Arrington, C., W. Hillison, and R. Jensen. "An Application of Analytical Hierarchy Process to Model Expert Judgments on Analytical Review Procedures." Journal of Accounting Research (Spring 1984): 298-312.
- Ashton, R. "Judgment Formation in the Evaluation of Internal Control: An Application of Brunswik's Lens Model." Ph.D. Dissertation. University of Minnesota, (June 1973).
- Ashton, R. "An Experimental Study of Internal Control Judgments." Journal of Accounting Research (Spring 1974): 143-157.
- Ashton, R., and P. Brown. "Descriptive Modeling of Auditors' Internal Control Judgments: Replication and Extension." Journal of Accounting Research (Spring 1980): 269-277.
- Ashton, R., and S. Kramer. "Students as Surrogates in Behavioral Accounting Research: Some Evidence." Journal of Accounting Research (Spring 1980): 1-15.
- Biggs, S., and J. Wild. "A Note of the Practice of Analytical Review." Auditing: A Journal of Practice & Theory (Spring 1984): 68-79.
- Biggs, S., and J. Wild. "An Investigation of Auditor Judgment in Analytical Review." The Accounting Review (October 1985): 607-633.
- Birnbaum, M. "Morality Judgment: Test of an Averaging Model with Differential Weights." Journal of Experimental Psychology (1973), 99: 395-399.
- Blocher, E., R. Esposito, and J. Willingham. "Auditors' Analytical Review Judgments for Payroll Expense." Auditing: A Journal of Practice & Theory (Fall 1983): 75-91.
- Bonner, S. "Experience Effects and Cue Choice in Analytical Risk Assessment." Working Paper, 1988.

- Brumfield, C., R. Elliott, and P. Jacobson. "Business Risk and the Audit Process." Journal of Accountancy (April 1983): 60-68.
- Cochran, W., and Cox, G. Experimental Designs. John Wiley & Sons, 1957.
- Cox, D. Planning of Experiments. John Wiley & Sons, 1958.
- Cushing, B., and J. Loebbecke. "Analytical Approaches to Audit Risk: A Survey and Analysis." Auditing: A Journal of Practice & Theory (Fall 1983): 23-41.
- Daroca, F., and W. Holder. "The Use of Analytical Procedures in Review and Audit Engagements." Auditing: A Journal of Practice & Theory (Spring 1985): 80-92.
- Dugan, M., J. Gentry, and K. Shriver. "The X-11 Model: A New Analytical Review Technique for the Auditor." Auditing: A Journal of Practice & Theory (Spring 1985): 11-22.
- Gaumnitz, B., T. Nunamaker, J. Surdick, and M. Thomas. "Auditor Consensus in Internal Control Evaluation and Audit Program Planning." Journal of Accounting Research (Autumn 1982): 745-755.
- Gibbins, M., and F. Wolf. "Auditors' Subjective Decision Environment -- The Case of a Normal External Audit." The Accounting Review (January 1982): 105-124.
- Grobstein, M., and P. Craig. "A Risk Analysis Approach to Auditing." Auditing: A Journal of Practice & Theory (Spring 1984): 1-16.
- Hamilton, R., and W. Wright. "Internal Control Judgments and Effects of Experience: Replications and Extensions." Journal of Accounting Research (Autumn 1982): 756-765.
- Hamilton, D., and M. Zanna. "Differential Weighting of Favorable and Unfavorable Attributes in Impressions of Personality." Journal of Experimental Research in Personality (1972), 6: 204-212.
- Hammond, K., T. Stewart, B. Brehmer, and D. Steinman. "Social Judgment Theory." In M. Kaplan and S. Schwartz (Eds.), Human Judgment and Decision Processes. Academic Press, Inc., 1975.
- Himmelfarb, S. "General Test of a Differential Weighted Averaging Model of Impression Formation." Journal of Experimental Social Psychology (1973), 9: 379-390.
- Hodges, B. "Effect of Valence on Relative Weighting in Impression Formation." Journal of Personality and Social Psychology (1974), 30: 378-381.
- Holder, W. "Analytical Review Procedures in Planning the Audit: An Application Study." Auditing: A Journal of Practice & Theory (Spring 1983): 100-107.

- Hylas, R., and R. Ashton. "Audit Detection of Financial Statement Errors." The Accounting Review (October 1982): 751-765.
- Joyce, E. "Expert Judgment in Audit Program Planning." Journal of Accounting Research (Supplement 1976): 29-59.
- Kaplan, S. "An Examination of the Effects of Environment and Explicit Internal Control Evaluation on Planned Audit Hours." Auditing: A Journal of Practice & Theory (Fall 1985): 12-25.
- Kaplan, S. and P. Reckers. "An Empirical Examination of Auditors' Initial Planning Processes." Auditing: A Journal of Practice & Theory (Fall 1984): 1-19.
- Keppel, G. Design and Analysis: A Researcher's Handbook. Prentice-Hall, Inc., 1982.
- Kinney, W. "ARIMA and Regression in Analytical Review: an Empirical Test." The Accounting Review (January 1978): 48-60.
- Kinney, W. "The Predictive Power of Limited Information in Preliminary Analytical Review: An Empirical Study." Journal of Accounting Research (Supplement 1979): 148-165.
- Kinney, W. "Attention-Directing Analytical Review Using Accounting Ratios: A Case Study." Auditing: A Journal of Practice & Theory (Spring 1987): 59-73.
- Kinney, W., and G. Salamon. "Regression Analysis in Auditing: A Comparison of Alternative Investigation Rules." Journal of Accounting Research (Autumn 1982): 350-366.
- Kinney, W., G. Salamon, and W. Uecker. Computer Assisted Analytical Review System. Accounting Education Series, No. 7. American Accounting Association, 1986.
- Kinney, W., and W. Uecker. "Mitigating the Consequences of Anchoring in Auditor Judgments." The Accounting Review (January 1982): 55-69.
- Krogstad, J., R. Ettenson, and J. Shanteau. "Context and Experience in Auditors' Materiality Judgments." Auditing: A Journal of Practice & Theory (Fall 1984): 54-73.
- Lampel, A., and N. Anderson. "Combining Visual and Verbal Information in an Impression-Formation Task." Journal of Personality and Social Psychology (1968), 9: 1-6.
- Lev, B. "On the Use of Index Models in Analytical Reviews by Auditors." Journal of Accounting Review (Autumn 1980): 524-550.
- Libby, R. "Availability and the Generation of Hypotheses in Analytical Review." Journal of Accounting Research (Autumn 1985): 648-667.

- Libby, R., J. Artman, and J. Willingham. "Process Susceptibility, Control Risk, and Audit Planning." The Accounting Review (April 1985): 212-230.
- Loebbecke, J., and P. Steinbart. "An Investigation of the Use of Preliminary Analytical Review to Provide Substantive Audit Evidence." Auditing: A Journal of Practice & Theory (Spring 1987): 74-89.
- Mock, T., and J. Turner. Internal Accounting Control Evaluation and Auditor Judgment, Audit Research Monograph No. 3. AICPA, 1981.
- Nanni, A. "An Exploration of the Mediating Effects of Auditor Experience and Position in Internal Accounting Control Evaluation." Accounting, Organizations and Society (1984): 149-163.
- Rosenbaum, M., and I. Levin. "Impression Formation as a Function of Source Credibility and Order of Presentation of Contradictory Information." Journal of Personality and Social Psychology (1968), 10: 167-174.
- Rosenbaum, M., and I. Levin. "Impression Formation as a Function of Source Credibility and the Polarity of Information." Journal of Personality and Social Psychology (1969), 12: 34-37.
- Sawyers, B., and N. Anderson. "Test of Integration Theory in Attitude Change." Journal of Personality and Social Psychology (1971), 18: 230-233.
- Shields, M., I. Solomon, and W. Waller. "Unaudited Book Values and the Accuracy of Auditors' Pre-sampling Judgments." Working Paper (August 1985).
- Srinidhi, B., and M. Vasarhelyi. "Auditor Judgment Concerning Establishment of Substantive Tests Based on Internal Control Reliability." Auditing: A Journal of Practice & Theory (Spring 1986): 64-76.
- Stringer, K. "A Statistical Technique for Analytical Review." Journal of Accounting Research (Supplement 1975): 1-9.
- Tabor, R., and J. Willis. "Empirical Evidence on the Changing Role of Analytical Review Procedures." Auditing: A Journal of Practice & Theory (Spring 1985): 93-109.
- Tucker, L. "A Suggested Alternative Formulation in the Developments by Hursch, Hammond, and Hursch, and by Hammond, Hursch, and Todd." Psychological Review (1964), 71: 528-530.
- Tversky, A., and D. Kahneman. "Judgment Under Uncertainty: Heuristics and Biases." Science (September 27, 1974): 1124-1131.

Wagenaar, W. "Note on the Construction of Digram-Balanced Latin Squares." Psychological Bulletin (1969), 72: 384-386.

Winer, B. Statistical Principles in Experimental Design. McGraw-Hill, 1962.

APPENDICES

APPENDIX A

Booklet 1

Volatile External Environment Condition

APPENDIX A
Booklet 1
Volatile External Environment Condition

INFORMATION INTEGRATION
IN AUDIT PLANNING

Booklet #1
Instructions and Background Information

INFORMATION INTEGRATION
IN AUDIT PLANNING

Professional audit judgment is the key component in audit planning. Publications of the AICPA and firms' audit manuals provide general guidelines on the planning of audits, but the actual planning of individual audits is left to the auditor's professional judgment. Very little is known about the actual judgment processes involved in audit planning. For instance, while the accounting literature suggests that certain types of information should be considered when planning an audit, the judgment processes involved in the use of this information when making decisions is not well understood.

The purpose of this study is to increase our understanding of how auditors use various types of information in audit planning. You will be asked to respond to several different cases that you may be faced with in a hypothetical audit. Your responses to these cases will allow us to gain insight into the manner in which you use the particular information presented in the cases when making audit planning decisions. The design of this study is based on techniques which have proved valuable for studying the judgment processes of members of other professions (e.g., doctors and security analysts).

In order to keep the time required for you to complete this study down to a reasonable level, and due to the nature of experimental studies such as this, we have not attempted to provide you with all of the information you might find useful when making audit planning decisions. However, in actual audit situations it is not likely that you are often able to have all the information at your disposal that you might like when planning an audit.

Please keep track of the time you spend in completing the cases, as the debriefing questionnaire at the end of booklet #3 will ask for this, and other, information. We estimate that it will take no longer than one hour of your time to complete the cases and the debriefing questionnaire.

It is important to the validity of this experiment that you do not confer with others when responding to the cases. Therefore, please work independently. If you desire, you may disassemble, highlight, or mark the booklets in any way you find useful.

When you have completed the cases and the debriefing questionnaire, please place the booklets in the enclosed envelope, seal it, and return it to Frank Burdine. Frank will then forward your envelope, unopened, to Michigan State University. Your responses to the cases will be kept strictly confidential.

Thank you very much for your cooperation and willingness to participate in this study. Your prompt completion and return of these booklets will be greatly appreciated.

INSTRUCTIONS

Assume you are the senior in charge of the year-end audit of the Casa-Carver Company. Casa-Carver has been an audit client of your firm for the past three years, but this is the first year that you have been in charge of the field work. You will have one junior-level staff accountant to assist you in this engagement. He has one year of auditing experience and has generally received high scores on his staff evaluations.

Casa-Carver is a privately held company. During 1988, Casa-Carver's total sales were approximately \$3,000,000. The purpose of this engagement is to issue an audit report on Casa-Carver's December 31, 1988 financial statements in order for Casa-Carver to comply with the terms of a 10-year bank loan for \$100,000 granted in 1986. The terms of the loan require Casa-Carver to provide the bank with audited financial statements annually, and to maintain a current ratio of 2.5:1. Your firm has given Casa-Carver an unqualified opinion in each of the previous two years.

In this experiment, the interest will be centered around how you would plan the audit of Casa-Carver's accounts receivable. You will be presented with several different situations. We are interested in the number of hours you would budget for the performance of substantive tests of details of the balance in Casa-Carver's accounts receivable for each situation.

The remaining pages of this booklet will provide you with certain information about Casa-Carver which the auditing literature suggests may be useful to you in this type of audit planning. The information on pages 5 through 8 of this booklet is background information about Casa-Carver and the industry in which it operates. This information will be held constant; i.e., it will not change from case to case.

In Booklet #2, you will be presented with additional information about Casa-Carver Company which will change from case to case. The manner in which this additional information is varied creates 10 different hypothetical circumstances you might be faced with as you plan the extent of substantive tests of details of the balance in Casa-Carver's accounts receivable. An example of a case you might be presented with is shown in Figure 1 on page 4 of this booklet.

For each case in Booklet #2, you will be asked to indicate the number of hours you would budget for substantive tests of details of the balance in Casa-Carver's Accounts Receivable account. As you respond to each situation, you should consider the information presented in only that particular case, along with the background information presented on pages 5 - 8 of this Booklet. DO NOT consider any of the information presented in previous cases.

In Booklet #3, you will be presented with several shorter cases than those in Booklet #2. Specifically, you will be presented with only information about Casa-Carver's internal control structure. This information will change from case to case. In this phase of the study, we are interested in how changes in Casa-Carver's internal control structure affect your evaluation of that structure. Accordingly, for each case in Booklet #3, you will be asked to evaluate the strength of Casa-Carver's internal control structure.

FIGURE 1: Sample Case

1. Last year's audit of receivables revealed NO material misstatements in the accounts. You believe the competence of the employees who deal with receivables to be HIGH, due in part to a low employee turnover rate.
2. Management has, in general, been quite COOPERATIVE in their dealings with your firm. Furthermore, management tends to follow CONSERVATIVE accounting practices.
3. Responses to four additional internal control questions (in addition to those presented in the internal control questionnaire in Booklet #1) are:

Billing and Recording: YES NO

Is billing performed by personnel independent of Credit and Cash Receipts personnel? X

Are customers' monthly statements reviewed and mailed by personnel independent of those who maintain the accounts receivable subsidiary ledger? X

Are write-offs of receivables approved by personnel independent of Cash Receipts and the credit manager? X

Is the accounts receivable subsidiary ledger reconciled with the general ledger at least monthly by the client? X

4. Your preliminary analytical review relating to receivables revealed the following:

<u>Bad Debt Expense</u>	No Significant Change
Gross Sales	from Last Year
<u>Allowance for Bad Debts</u>	No Significant Change
Accounts Receivable	from Last Year
Average Number of Days that	No Significant Change
Accounts Receivable Were Outstanding	from Last Year
Comparison of Individual Customer's	3 out of 55 Accounts
Balances over \$3,000 with Previous	Changed by More Than
Year's Balances	20%.

BACKGROUND INFORMATION

Casa-Carver Company began operations 3 years ago. The company designs, develops, makes and distributes high-fidelity audio compact disc players and specialized circuits for home entertainment use. Products are sold throughout the United States through 400 retail outlets.

The audio compact disc player industry appears to be highly volatile. At the end of 1987, there were 70 firms in the business of manufacturing audio compact disc players. During 1988, 7 (10%) of those firms withdrew from the compact disc player industry, while 23 firms entered the compact disc player industry. Thus, as of December 31, 1988, there were 86 firms manufacturing compact disc players, a net increase of 23% over last year. Rapid obsolescence of current models of players is a major concern for Casa-Carver.

Profits for Casa-Carver over the past three years have been IRREGULAR, with a declining trend.

You have just been informed that Casa-Carver is planning to go public within the next year.

Background Information is
Continued on Next Page

Customers. Casa-Carver sells its products throughout the United States through 400 retail outlets. Casa-Carver's books report an unaudited accounts receivable balance at the end of 1988 of \$425,400.

You have received the following listing of customer's accounts from Casa-Carver's controller:

<u>Number of Customers</u>	<u>Unadjusted Balance in Customers' Receivables (12-31-88)</u>	<u>Total</u>
0	\$ 20,000 - \$ 50,000	\$ -0-
1	15,000 - 19,999	16,000
1	10,000 - 14,999	12,000
3	5,000 - 9,999	18,000
50	3,000 - 4,999	175,000
42	2,000 - 2,999	105,000
33	1,000 - 1,999	39,600
36	500 - 999	21,600
74	250 - 499	22,200
<u>160</u>	0 - 249	<u>16,000</u>
<u>400</u>		<u>\$425,400</u>

Internal Controls Over Receivables. As part of your review and testing of Casa-Carver's internal control system over receivables, your assistant has prepared the internal control questionnaire on page 7. Casa-Carver's accounting for accounts receivable is done manually.

Further information about Casa-Carver's controls will be presented in the individual cases presented in Booklet #2. The individual cases will also provide you with other information, including information about the results of the previous year's audit, the accounting practices used by management, the competence of the employees who deal with receivables, and the relationship between management and you, the auditors.

INTERNAL CONTROL QUESTIONNAIRE: ACCOUNTS RECEIVABLEName of Client: Casa-Carver Company For Period Ended: 12-31-88Customer Order:YES NO

- | | |
|--|---|
| 1. Are policies and procedures for accepting and approving customer orders clearly defined? | X |
| 2. Are prenumbered sales orders prepared for all approved customers orders? | X |
| 3. Is current information regarding prices, policies on discounts, freight, and returned goods available and communicated to Customer Order personnel? | X |
| 4. Are copies of sales orders forwarded to Shipping and Billing? | X |

Credit:

- | | |
|--|---|
| 1. Are Credit personnel independent of Accounting, Billing, and Cash Collection personnel? | X |
| 2. Is credit investigated before approval? | X |
| 3. Is information about past due accounts communicated to Credit personnel? | X |

Shipping:

- | | |
|--|---|
| 1. Are goods shipped only in accordance with approved sales orders? | X |
| 2. Are shipping documents numerically accounted for to assure that all shipped goods are billed? | X |
| 3. Are Shipping personnel independent of Billing, Cash Collection, and Accounting personnel? | X |

Billing and Recording:

- | | | |
|---|---|---|
| 1. Are accounts receivable confirmed by the client during the year? | | X |
| 2. Are sales invoices matched with approved sales orders and shipping documents, and checked for clerical accuracy? | X | |
| 3. Are prenumbered sales invoices prepared for all shipped goods? | X | |

NOTE: Each case in Booklet #2 will present you with four additional items of information about Casa-Carver's internal control system over receivables.

Financial Information --**Balance Sheet December 31, 1988:**

Assets--	
Cash	\$ 80,900
Marketable Securities	43,300
Accounts Receivable	425,400
Less: Allowance for Bad Debts	- 29,800
Inventories	440,560
Prepayments	67,100
Total Current Assets	1,027,460
Net Property	792,240
Other Assets	10,600
Total Assets	\$ 1,830,300

Equities--	
Accounts Payable	324,000
Accruals	82,100
Total Current Liabilities	406,100
Notes Payable	150,000
Total Liabilities	556,100

Owner's Equity 1,274,200

Total Liabilities and
Owner's Equity \$ 1,830,300

Earnings for 1988:

Net Sales	\$ 2,493,400
Income Before Taxes	102,600
Income Taxes	25,600
Net Income	77,000

Employees, December 31, 1988: 64

This completes the unchanging background information. It is important that you base your budgeted time estimates in Booklet #2 only on the data presented in this experiment. A careful effort has been made to maintain the realism of the information provided to you. While Casa-Carver is a fictional company, IT HAS BEEN MODELED AFTER AN ACTUAL LYNNWOOD, WASHINGTON MANUFACTURER OF AUDIO COMPACT DISC PLAYERS.

BEFORE TURNING TO THE CASES IN BOOKLET #2, PLEASE
COMPLETE THE EIGHT ITEMS ON THE FOLLOWING PAGES:

1. Based on the background information presented above, please indicate the number of hours you would budget for this year's audit of accounts receivable:

_____ Hours

2. Based on the background information presented above, PLEASE INDICATE YOUR ACCEPTABLE LEVEL OF OVERALL RISK for this year's audit of Casa-Carver Company by placing a slash mark on the scale below:

(The Acceptable Level of Overall Risk is the risk that you are willing to take that the financial statements may be materially misstated after the audit is completed and an unqualified opinion has been reached).

Low	_____	High
(E.g., 1%)		(E.g., 9%)

On pages 11 - 15 are partial examples of additional information you might be presented with in Booklet #2. Please read each item and provide the information requested.

3. Last year's audit of receivables revealed NO material misstatements in the accounts. You believe the competence of the employees who deal with receivables to be HIGH, due in part to a low employee turnover rate.

Management has, in general, been quite COOPERATIVE in their dealings with your firm. Furthermore, management tends to follow CONSERVATIVE accounting practices.

Based only on the additional information above, WHAT DO YOU BELIEVE IS THE LIKELIHOOD THAT MATERIAL ERRORS OR IRREGULARITIES HAVE OCCURRED in Casa-Carver's accounts receivable during 1988?

Not	_____	100%
Likely		Likelihood
At All		

4. Last year's audit of receivables revealed material misstatements in approximately 20% of the accounts. You believe the competence of the employees who deal with receivables to be rather LOW, due in part to a high employee turnover rate.

Management has, in general, been rather UNCOOPERATIVE in their dealings with your firm. Furthermore, management has been quite aggressive in trying to improve the appearance of their company's financial position by following UNCONSERVATIVE accounting practices.

Based only on the additional information above, WHAT DO YOU BELIEVE IS THE LIKELIHOOD THAT MATERIAL ERRORS OR IRREGULARITIES HAVE OCCURRED in Casa-Carver's accounts receivable during 1988?

Not	_____	100%
Likely		Likelihood
At All		

5. Responses to four additional internal control questions (in addition to those presented in the internal control questionnaire on page 7) are:

<u>Billing and Recording:</u>	<u>YES</u>	<u>NO</u>
Is billing performed by personnel independent of Credit and Cash Receipts personnel?	X	
Are customers' monthly statements reviewed and mailed by personnel independent of those who maintain the accounts receivable subsidiary ledger?	X	
Are write-offs of receivables approved by personnel independent of Cash Receipts and the credit manager?	X	
Is the accounts receivable subsidiary ledger reconciled with the general ledger at least monthly by the client?	X	

Please indicate how much you agree with the following statement by placing a slash mark on the scale below:

The responses on the internal control questionnaire (the four questions above and those on page 7) suggest that Casa-Carver's internal control structure is HIGHLY EFFECTIVE AT PREVENTING AND/OR DETECTING ERRORS relating to accounts receivable.

Strongly
Disagree

Strongly
Agree

6. Responses to four additional internal control questions (in addition to those presented in the internal control questionnaire on page 7) are:

<u>Billing and Recording:</u>	<u>YES</u>	<u>NO</u>
Is billing performed by personnel independent of Credit and Cash Receipts personnel?		X
Are customers' monthly statements reviewed and mailed by personnel independent of those who maintain the accounts receivable subsidiary ledger?		X
Are write-offs of receivables approved by personnel independent of Cash Receipts and the credit manager?		X
Is the accounts receivable subsidiary ledger reconciled with the general ledger at least monthly by the client?		X

Please indicate how much you agree with the following statement by placing a slash mark on the scale below:

The responses on the internal control questionnaire (the four questions above and those on page 7) suggest that Casa-Carver's internal control structure is HIGHLY EFFECTIVE AT PREVENTING AND/OR DETECTING ERRORS relating to accounts receivable.

Strongly
Disagree

Strongly
Agree

7. Your preliminary analytical review relating to receivables revealed the following:

<u>Bad Debt Expense</u> Gross Sales	No Significant Change from Last Year
<u>Allowance for Bad Debts</u> Accounts Receivable	No Significant Change from Last Year
Average Number of Days that Accounts Receivable Were Outstanding	No Significant Change from Last Year
Comparison of Individual Customer's Balances over \$3,000 with Previous Year's Balances	3 out of 55 Accounts Changed by More Than 20%.

Please indicate how much you agree with the following statement by placing a slash mark on the scale below:

The results of the analytical procedures above suggest there is a HIGH
LIKELIHOOD that the accounts receivable balance may be MATERIALLY
MISSTATED.

Strongly
Disagree

Strongly
Agree

8. Your preliminary analytical review relating to receivables revealed the following:

<u>Bad Debt Expense</u> Gross Sales	Significant Increase from Last Year
<u>Allowance for Bad Debts</u> Accounts Receivable	Significant Decrease from Last Year
Average Number of Days that Accounts Receivable Were Outstanding	Significant Increase from Last Year
Comparison of Individual Customer's Balances over \$3,000 with Previous Year's Balances	21 out of 55 Accounts Changed by More Than 20%.

Client management is unable to explain why these changes occurred.

Please indicate how much you agree with the following statement by placing a slash mark on the scale below:

The results of the analytical procedures above suggest there is a HIGH LIKELIHOOD that the accounts receivable balance may be MATERIALLY MISSTATED.

Strongly
Disagree

Strongly
Agree

Thank You. Please turn now to Booklet #2.

Sub _____
Ord _____
Ind _____

APPENDIX B

Booklet 2 (Partial)

Sample Audit Planning Case

APPENDIX B
Booklet 2 (Partial)
Sample Audit Planning Case

INFORMATION INTEGRATION
IN AUDIT PLANNING

Booklet #2
Audit Planning Cases

INSTRUCTIONS

For each of the following cases in this booklet, please indicate the number of hours you would budget for the performance of substantive tests of details of the balance in Casa-Carver's accounts receivable.

To help you with your planning, keep in mind that you may need to include time for the following substantive procedures:

-
- Confirmation of Receivables.
 - Examination of the appropriateness of the balance in the Allowance for Doubtful Accounts.
 - Investigation of accounts written off as uncollectible.
 - Review of cash collected subsequent to year-end from receivables.
 - Testing of year-end sales cutoff.
 - Other relevant tests of details of the balance in accounts receivable that you consider necessary.
-

In Cases #1 & #2, you will be given only two additional pieces of information. This information will pertain to the results of the previous year's audit, and your relationship with Casa-Carver's management. Please base your responses to these two cases on ONLY the information provided in each case, and on the background information in Booklet #1. However, when considering the background information in Booklet #1, please IGNORE the internal control questionnaire on page 7. Assume that internal controls have NOT yet been reviewed, and that analytical procedures have NOT yet been performed.

In Cases #3 - #10, you will be presented with four items of information. This information will be similar to the information provided in the first two cases, but will also include details about Casa-Carver's internal controls and the results of analytical procedures.

Case _____

1. Last year's audit of receivables revealed NO material misstatements in the accounts. You believe the competence of the employees who deal with receivables to be HIGH, due in part to a low employee turnover rate.
2. Management has, in general, been quite COOPERATIVE in their dealings with your firm. Furthermore, management tends to follow CONSERVATIVE accounting practices.
3. Responses to four additional internal control questions (in addition to those presented in the internal control questionnaire in Booklet #1) are:

Billing and Recording:YES NO

Is billing performed by personnel independent of Credit and Cash Receipts personnel?

X

Are customers' monthly statements reviewed and mailed by personnel independent of those who maintain the accounts receivable subsidiary ledger?

X

Are write-offs of receivables approved by personnel independent of Cash Receipts and the credit manager?

X

Is the accounts receivable subsidiary ledger reconciled with the general ledger at least monthly by the client?

X

4. Your preliminary analytical review relating to receivables revealed the following:

Bad Debt Expense

Gross Sales

No Significant Change
from Last YearAllowance for Bad Debts

Accounts Receivable

No Significant Change
from Last Year

Average Number of Days that Accounts Receivable Were Outstanding

No Significant Change
from Last Year

Comparison of Individual Customer's Balances over \$3,000 with Previous Year's Balances

3 out of 55 Accounts
Changed by More Than
20%.

Note: Client management is unable to explain why these changes occurred.

Given the information on the preceding page, and the background information in Booklet #1, please INDICATE HOW MUCH TIME YOU WOULD PLAN for this year's substantive tests of details of the accounts receivable balance.

TOTAL NUMBER OF HOURS planned for this year's audit of accounts receivable is:

_____ HOURS

In addition, would you please provide a second INDICATION OF YOUR PLANNED EXTENT OF TESTING of accounts receivable by placing a slash on the scale below:

No
Testing
of Receivables

Extensive
Testing
of Receivables

APPENDIX C

Booklet 3 (Partial)

(Sample Internal Control Evaluation Case
and Debriefing Questionnaire)

APPENDIX C
Booklet 3 (Partial)
(Sample Internal Control Evaluation Case
and Debriefing Questionnaire)

INFORMATION INTEGRATION
IN AUDIT PLANNING

Booklet #3
Internal Control Evaluation Cases

INSTRUCTIONS

In the following cases, you will be asked to evaluate the strength of Casa-Carver's internal control structure. In this phase of the experiment, we are interested in how changes in Casa-Carver's internal control structure affect your evaluations of that structure.

When responding to each of the following cases, you should consider the background information in booklet #1, along with the particular information presented in only that case. Each of the cases should be considered independently from one another. In addition, they should be considered independently from the cases in Booklet #2.

NOTE: In Cases #1 & #2, you will be provided with information on only two particular internal controls. In Cases #3 - #10, you will be provided with information on four particular controls.

Case _____

Billing and Recording:YESNO

Is billing performed by personnel independent of Credit and Cash Receipts personnel?

X

Are customers' monthly statements reviewed and mailed by personnel independent of those who maintain the accounts receivable subsidiary ledger?

X

Are write-offs of receivables approved by personnel independent of Cash Receipts and the credit manager?

X

Is the accounts receivable subsidiary ledger reconciled with the general ledger at least monthly by the client?

X

Based on the above information, and on the background information in Booklet #1, please indicate your assessment of the EFFECTIVENESS OF CASA-CARVER'S INTERNAL CONTROL STRUCTURE at preventing and detecting material misstatements in accounts receivable by placing a slash mark on the scale below.

Extremely
Weak

Adequate to
Strong

DEBRIEFING QUESTIONNAIRE

Please complete the following questions:

1. What was your major as an undergraduate student? (Circle one)
- | | |
|---|---|
| Accounting | 1 |
| Finance | 2 |
| Business (other than accounting or finance) . . | 3 |
| Other (please specify) _____ . . | 4 |

2. (a) Do you have a graduate degree? YES NO

- (b) If YES, what was your degree? (Circle one)
- | | |
|--|---|
| M.S. in Accounting | 1 |
| M.B.A. with emphasis in Accounting | 2 |
| M.B.A. with emphasis in area other than Accounting | 3 |
| Other (please specify) _____ . . | 4 |

3. How many undergraduate or graduate AUDITING courses have you completed?

_____ Courses

4. Please list any firm training seminars you have attended:

5. Please indicate where you are employed:

City _____, State _____

6. Please indicate your present level within the Firm:

(Circle one)

E	1
D1	2
D2	3
Senior.	4
Manager	5
Senior Manager.	6
Partner	7
Other (please describe) _____ .	8

How long have you been at this level?

_____ Years _____ Months

7. To the nearest year, how long have you been working as an auditor?

_____ Years

8. Are you a CPA? YES NO

9. In what year were you born? 19____.

10. Do you consider yourself to be an industry specialist?

(Circle one)

YES (What Industry?) _____ . .	1
NO	2

11. Have you ever audited a manufacturer of compact disc players? YES NO

12. Have you ever audited a manufacturer of lighting fixtures or ceiling fans? YES NO

13. Compared with other auditors, do you feel you tend to perform more or less extensive audit procedures?

(Circle one)

Much more extensive	1
Somewhat more extensive	2
About the same as most other auditors.	3
Somewhat less extensive	4
Much less extensive	5

14. Based upon your audit program planning experience, do you tend to underestimate or overestimate the time required to actually complete the work you have planned?

(Circle one)

I have no audit-planning experience	1
Usually overestimate	2
Usually reasonably accurate	3
Usually underestimate	4

15. Compared with other auditors, do you feel you tend to accept more or less risk?

(Circle one)

Much more willing to accept risk	1
Somewhat more willing to accept risk	2
About the same as most	3
Somewhat less willing to accept risk	4
Much less willing to accept risk	5

16. For what percentage of your audit engagements of manufacturing firms would you consider the client's system of internal controls to be:

Very Strong	_____ %
Fairly Strong	_____ %
Fairly Weak	_____ %
Very Weak	_____ %
	<u>100</u> %

17. Would you have found additional information helpful in completing the cases? YES NO

If YES, please specify. _____

18. How realistic did you find the experiment?

(Circle One)

Very Realistic	1
Somewhat Realistic.	2
Somewhat Unrealistic.	3
Very Unrealistic	4

19. How interesting did you find this experiment?

(Circle One)

Very interesting	1
Fairly Interesting	2
Fairly Boring	3
Very Boring	4

20. Would you like a copy of the results of this study? YES NO

If yes, please indicate where you would like it sent:

Name_____

Address_____

21. Additional comments, if any:

22. How long did it take you to complete this experiment (excluding any breaks you may have taken)?

_____ hour(s) _____ minutes

Thank you very much for your participation in this study. Please check to make sure that you have completed all of the cases in Booklets #2 and #3, and then return all booklets in the enclosed envelope to Frank Burdine, who will forward them to Michigan State University.

MICHIGAN STATE UNIV. LIBRARIES



31293006059608