



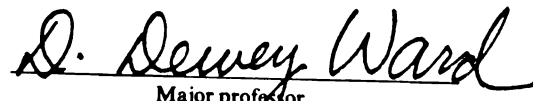
This is to certify that the
dissertation entitled

The Use of Working Paper Review in Reducing the Risk
of Type II Errors

presented by
Brian Ballou

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Accounting


Major professor

Date 7-9-96

LIBRARY

Michigan State University

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
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

**THE USE OF AUDIT WORKING PAPER REVIEW IN REDUCING THE RISK OF
TYPE II ERRORS**

By

Brian Ballou

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Accounting

1996

ABSTRACT

THE USE OF AUDIT WORKING PAPER REVIEW IN REDUCING THE RISK OF TYPE II ERRORS

By

Brian Ballou

This dissertation tested whether the audit review process for working papers can help reduce the risk of Type II errors by offsetting the decision making biases encountered when performing analytical procedures. A Type II error occurs when an auditor incorrectly accepts a client explanation indicating that no error or irregularity has occurred, when, in fact, one or the other has happened. The impact of two external cues, the instruction to write review notes and the provision of a cognitive decision aid to enhance memory retrieval, is examined between groups of audit reviewers with specialized knowledge in the financial services industry.

The results indicate that the decision aid reduces the risk of Type II errors; however, the instruction to provide review notes increases the risk of Type II errors. Using Shiffrin and Schneider's (1977) information processing model and Kahneman's (1977) capacity model of attention, these effects might have occurred because of the reviewers' limited attentional capacity, which is needed in order to consider the different aspects of the review process. Reviewers who performed the review note process might have experienced an excessive strain on capacity that increased the risk of committing Type II errors. Reviewers who were provided a decision aid might have experienced capacity relief that decreased the risk of committing Type II errors. These effects were

offset when reviewers are exposed to both a cognitive decision aid and instruction to provide written review notes.

Also, the manipulations had a stronger impact on reviewers when their responses indicated that they were more likely to commit Type II errors *a priori*. These reviewers were more concerned with performing costly, additional procedures and other aspect of the review besides Type II errors (documentation issues, cohesiveness of working papers, etc.).

ACKNOWLEDGMENTS

First, I express gratitude to my dissertation committee-- Dr. D. Dewey Ward (Chair), Dr. Frank Boster, Dr. Susan Haka, and Dr. Joan Luft---for their support and insights throughout the dissertation process. I also am particularly thankful to Dr. Alvin Arens for his contributions to the dissertation, which included the initial contact with both international accounting firms that participated in the experiment. In addition, I am grateful to Drs. Harold Sollenberger, Matthew Anderson, Sevrin Grabski, and Kathy Petroni for their support, mentoring, and teaching during the doctoral program.

I wish to thank both international accounting firms for their participation in this project. In particular, I am grateful to the partner group that helped design the working papers and background information and select the specialized industry to use for the experiment. Also, I thank the participants in the experiment.

I am particularly grateful for my parents, Richard and Janice Ballou, for their emotional and financial support throughout my entire life. I have been blessed with loving role models. The calls and letters from friends and other family members are also appreciated.

Finally, and most importantly, I wish to thank my wife, Maria, and our sons, Alex and Grant. Their love, patience, and encouragement provided the motivation for me to complete this dissertation and continue with a career in academia.

TABLE OF CONTENTS

LIST OF TABLES	vii
LIST OF FIGURES	viii
1. INTRODUCTION	1
1.1 Motivation and Review of Working Paper Review Literature	9
1.2 Overview of Hypotheses Examined	16
1.3 Overview of Research Design and Tests	22
1.4 Organization of the Dissertation	23
2. THEORY AND HYPOTHESES DEVELOPMENT	25
2.1 Theories and Literature Relevant for Effective Working Paper Review	25
2.1.1 Cognitive Effort: Attention Capacity and Motivation	26
2.1.2 Alternative Hypothesis Generation and Evaluation	32
2.2 Development of Hypotheses	35
2.2.1 The Impact of Review Notes	35
2.2.2 A Retrieval Plan to Relieve Demands on Capacity	41
2.2.3 The Combination of Review Notes and a Decision Aid	47
2.3 Task and Environmental Considerations	52
2.3.1 Belief Adjustment for Inherited Hypotheses	52
2.3.2 <i>A Priori</i> Likelihood of Committing Type II Errors	54
2.4 Summary of Chapter	59
3. RESEARCH DESIGN	61
3.1 Subjects	61
3.2 Firms	65
3.3 Task Description and Administration of the Experiment	67
3.4 Dependent Construct and Measures	79
3.5 Additional Data Gathered	86
3.5.1 Reviewer Confidence	87
3.5.2 Reviewer Concern for Efficiency and Effectiveness	87
3.5.3 Reviewer Perception of Client and Preparer	89
3.5.4 Manipulation Checks	90
3.5.5 Measures of Reviewer's Professional Experience	90
3.5.6 Reviewer Effort During Experiment	91
3.6 Summary of Chapter	92

4.	EXPERIMENTAL TESTS AND RESULTS	93
4.1	Specification and Measurement of Dependent Construct	93
4.2	Classification of Review Notes	99
4.3	Tests of Hypotheses	102
4.4	Tests for Environmental and Task-Specific Variables	110
4.4.1	<i>A Priori</i> Likelihood of Committing Type II Errors	111
4.4.2	Reviewer Motivation	114
4.5	Prior Influence and Concomitant Variables	119
4.5.1	Reviewer Confidence	119
4.5.2	Measures of Reviewers' Professional Experience	120
4.6	Manipulation Checks	122
4.7	Analysis of Alternative Explanation Listings	123
4.8	Summary of Chapter	126
5.	CONCLUSION	129
5.1	Summary of Study	130
5.2	Summary of Hypotheses and Results	136
5.3	Contributions	143
5.4	Limitations	146
5.5	Suggestions for Future Research	147
5.6	Concluding Statement	149
	APPENDIX	151
	LIST OF REFERENCES	161

LIST OF TABLES

Table 1 - Specification and Descriptive Statistics for Dependent Construct	98
Table 2 - Impact of Review Notes and/or Decision Aid on Review Judgments	105
Table 3 - Contrast Coding Analysis for “Magic Cell” Model	109
Table 4 - Tests for <i>A Priori</i> Likelihood of Committing Type II Errors	115
Table 5 - Comparison of Conscious Processing Reported by Reviewers	118
Table 6 - Professional Experience Descriptive Data	121

LIST OF FIGURES

Figure 1 - Kennedy's (1995) Debiasing Framework	27
Figure 2 - Decision Aid Designed to Relieve Demands on Capacity	45
Figure 3 - Summary of Hypotheses	51
Figure 4 - Research Design	68
Figure 5 - Cover Page of Experiment Given to Reviewers	71
Figure 6 - Instructions Given To Reviewers for Experiment	72
Figure 7 - Background Information on Client and Engagement Provided to Reviewers	74
Figure 8 - Instructions to Provide Review Notes Given to Reviewers	75
Figure 9 - Analytical Procedure Working Paper Provided to Subjects for Review Panel A	76
Panel B	77
Figure 10 - Questions that Measure Reviewer's Judgments of Working Paper Conclusions	82
Figure 11 - Hypothesized Relationships Between Dependent and Independent Variables Panel A: Review Notes Cue Counterexplanation/Alternatives	84
Panel B: Review Notes <i>Do Not</i> Cue Counterexplanation/Alternatives	85
Figure 12 - Diagram of Confirmatory Factor Analysis	97
Figure 13 - Plot of Means for Dependent Measure	100

Chapter One

INTRODUCTION

The audit review process is a key quality control function of any audit engagement and is required by generally accepted auditing standards (AICPA 1972, 1978). The audit review process consists of the evaluation of working papers that have been prepared and signed-off as complete by a subordinate auditor (preparer). The first review usually is performed by either the preparer's immediate supervisor or the supervisor at the next highest level (reviewer). Because most audit procedures, including many analytical procedures, generally are performed by staff auditors, the first review is often performed by a senior auditor or a manager. Also, the first review is performed at a detailed level, with additional reviews most likely performed at a more general level. The reviewer acknowledges by documenting approval (i.e., signing off) on the working paper that the procedures performed by the preparer are prepared sufficiently and in accordance with the approved audit program. Thus, the working paper should be defensible and justifiable if reviewed by another party, such as an engagement partner, a second partner, an external peer, or a court of law (Davis and Solomon 1989).

This dissertation examines the effectiveness of the working paper review process as a quality control mechanism to reduce the probability that a material misstatement¹ of the financial statements is not detected during the audit. The role of the review process

¹ In this study, the term "misstatement" refers only to errors or irregularities that have a material impact on the financial statements.

examined in this study is to help auditors ensure that nonmisstatement explanations from clients are not accepted when misstatements actually have occurred.²

Viewing the review process as a mechanism to reduce the risk of Type II errors is consistent with the reasoning stated by the AICPA Auditing Standards Board in Statement on Auditing Standard (SAS) No. 22, Planning and Supervision (AICPA 1978). SAS No. 22 states that “the work performed by each assistant should be reviewed to determine whether it was adequately performed and to evaluate whether the results are consistent with the conclusions to be presented in the auditor's report.”

To help answer the research question, an audit task that relies heavily on the review process as a means of quality control was selected. The task, substantive analytical procedures, involves auditor comparisons of quantifiable data to expectations. When material fluctuations from expectations occur, the auditor must investigate the fluctuation and gather evidence to support the conclusion. Auditors are required to perform analytical procedures during the planning and completion phase of the audit (AICPA 1988). However, these tests help direct the auditor's attention of where to gather evidence during the substantive phase of the audit, when evidence is gathered supporting detailed transactions and year-end account balances. Substantive analytical procedures involve the use of analytical procedures as primary sources of evidence for account

² A nonmisstatement explanation is defined as an explanation from the client that no intentional or unintentional misstatements occurred to cause the unexpected fluctuation in the account being examined.

balances. They are not required but provide auditors with an efficient way to gather evidence (Arens and Loebecke 1994). Ameen and Strawser (1994) found that auditors are devoting a significant amount of time during substantive testing to analytical procedures, as high as 46% of the time for Big Six firms.

Substantive analytical procedures are important to use as a context within which to study the working paper review process because (1) they are used as a primary source of evidence for account balances and (2) they involve the auditor's ability to test the client's hypothesis for why an unexpected fluctuation occurred. The use of these procedures as primary evidences means that auditors rely on the results of the analytical procedures when concluding whether the accounts in question are stated fairly. The auditor's ability to test the client's hypothesis is relevant because should the auditor accept a nonmisstatement explanation from the client when a misstatement actually happened, a Type II error will have occurred and the audit firm is exposed should users incur losses. These errors can prove costly to auditors, if failure to detect a misstatement results in litigation brought against the auditor by users of the financial statements (Palmrose 1991).

The use of working paper review for substantive analytical procedures is a key quality control mechanism in part because extant research has demonstrated deficiencies in preparer judgments when performing analytical procedures. A general finding in these studies is that preparers do not counterexplain (develop reasons why a preparer's conclusion might not be true) or generate and evaluate alternative hypotheses (other reasons that explain the fluctuation) when performing analytical procedures without

explicitly being instructed to do so (cf. Heiman 1990, Bedard and Biggs 1991a and 1991b, and Koonce 1992). As mentioned previously, this failure can be costly because the analytical procedures are being used as primary evidence and misstatements may go undetected.

Failure to appropriately consider and rule out alternative hypotheses of misstatements when the client provides a nonmisstatement explanation substantially increases the risk of Type II errors. Research has demonstrated that preparers focus primarily on seeking confirmatory evidence (see Church 1990 for a review in auditing). Therefore, reviewers are needed to challenge whether preparers have gathered enough evidence to rule out misstatement explanations. Thus, a primary objective of the review process is to minimize the risk of Type II errors by evaluating the extent to which relevant alternative hypotheses have been considered and ruled out by the evidence documented in the working papers.

Because the working paper review process is an explicit quality control function, the organization of the task and differences in knowledge should better enable a reviewer than a preparer to consider alternative hypotheses for analytical procedure results. Reviewers are trained to ensure that preparers' results are consistent with the conclusions to be presented in the auditor's report (AICPA 1978). Libby and Trotman (1993) provide some evidence for this claim by demonstrating that reviewers who evaluated the information used for preparers' initial conclusions recalled more information that was inconsistent than consistent with the preparers' conclusions. Furthermore, Frederick (1991) and Libby and Frederick (1990) demonstrate that those auditors who perform

detailed working paper review, seniors and managers, have more developed knowledge than more inexperienced auditors of transaction cycles and alternative hypotheses.

However, it is possible that the number of considerations involved in the working paper review process for analytical procedures can make it difficult for reviewers to consider alternative hypotheses for cognitive reasons. In general, there is agreement in cognitive psychology that individuals have limited attentional capacity, which places a constraint on the access to stored knowledge at any given point in time (Ashcraft 1994). As part of the working paper review process for analytical procedures, a reviewer must evaluate the preparer's hypothesis, quality and extent of evidence gathered, relevance of evidence gathered, organization and preparation of the working papers, and relationship of the preparer's conclusions to other findings in the audit. In addition, reviewers must evaluate the risks associated with receiving evidence from two persuasive communicators, the client and the preparer (Rich, Solomon, and Trotman 1994). These risks involve assessments by the reviewer of the extent to which the client and preparer are competent and reporting truthfully (Rich et al. 1994, Eagly and Chaiken 1993).

It is reasonable based on extant research (c.f. Libby 1985 and Libby and Frederick 1990) to assume that the experienced auditors who perform working paper reviews have the knowledge to consider alternative misstatement hypotheses. The question remains of whether the review process is affected by the difficulty that the reviewer may have considering all aspects of the review at the same time. It is possible that reviewers are not effective at internally organizing the different considerations. Such inability can impede information processing and long-term memory access (Shiffrin and Schneider 1977).

The difficulty for reviewers making the number of other considerations as part of the working paper review can be demonstrated with the following description. Reviewers must review the working papers to ensure that they are legible, coherent, and mechanically accurate. Also, reviewers analyze the working papers to determine whether the information contained within the papers is internally consistent. They are also concerned with how the working papers in question fit into the cycle being tested and the audit as a whole. Furthermore, reviewers must address the extent to which they believe the client or the preparer is knowledgeable about the information contained in the working papers and motivated not to report truthfully. Thus, a reviewer might have difficulty considering alternative misstatement hypotheses without being cued because of these demands on their attentional capacity.

In terms of Kennedy's (1993, 1995) debiasing framework, tests are needed to see whether the review process creates excess demands on cognitive effort or aids internal data (stored knowledge) retrieval to perform the review process for analytical procedures. Specifically, it is important to understand better the extent that reviewers counterexplain or generate and evaluate alternative hypotheses to preparers' conclusions. In this study, two external cues are introduced to help gain a better understanding of the cognitive demands placed on the reviewer when examining working papers for substantive analytical procedures.

The first cue is the instruction to provide written review notes and is included because it represents current practice in which reviewers currently write review notes to

document working paper deficiencies.³ These review notes provide preparers with explicit instructions on what follow-up procedures need to be performed and reviewers with a documented audit trail for follow-up review procedures. Instruction to provide review notes may cue counterexplanation and aid the consideration of alternative hypotheses. Thus, the review notes themselves may proxy for counterexplanation. However, if instruction to provide review notes does not cue counterexplanation, the requirement to state explicitly any noted deficiencies and instruct the preparer on a course of action could increase the amount of cognitive effort, defined as available attentional capacity or motivation (Kennedy 1995, 1993), needed for the review process. Thus, the process of considering and writing review notes might place an added strain on attentional capacity, which is limited, or reduce the motivation of the reviewer. This study will control for changes in motivation to help identify which component of cognitive effort is affected by instruction to provide review notes.

The second cue is the provision of a cognitive decision aid containing general audit-firm guidance on how to evaluate analytical procedures. The cognitive decision aid is designed to create a retrieval plan (Medin and Ross 1992) for reviewers, which should help them internally organize the considerations for analytical procedures. The improved organization should free up limited attentional capacity for more consideration of

³ These review notes are not included in the working papers. Their sole purpose is to communicate follow-up procedures from reviewer to preparer. Due to liability concerns, these notes are kept only until the end of the audit and then destroyed.

alternative misstatement explanations. Informal, unstructured guidance like that used for this retrieval plan has been shown to enhance performance of analytical procedures by preparers (McDaniel and Kinney 1995; McDaniel 1990; and Bonner, Nelson, and Libby 1996).

To help answer whether the current review process is effective and/or can be improved, reviewers will be exposed to either one, both, or neither external cue in a between-subjects experiment. If instruction to provide review notes cues counterexplanation, then the current review process should help reduce the likelihood of committing Type II errors. Furthermore the introduction of a cognitive decision aid should further decrease the risk of committing such errors. Conversely, if instruction to provide review notes does not cue counterexplanation, then the current review process should increase the likelihood of committing Type II errors compared to reviews that do not include review notes. However, the introduction of a cognitive decision aid should help offset this difficulty by reducing the likelihood of committing Type II errors.

Prior research has not addressed the question of whether the working paper review process is effective for considering alternative explanations or could be improved with external guidance. The results of this experiment can help identify the working paper review process as an important but cognitively challenging quality control mechanism for substantive analytical procedures. The results are particularly important to consider based on the heavy use of simple, judgmental analytical procedures by auditing firms

(Ameen and Strawser 1994).⁴ Based on discussions with several firms, this question could be even more important as firms move to more “risk-based” audit approaches, which include heavier reliance on substantive analytical procedures for all accounts that the firms do not consider to be “high risk” for any given client or year.

The remainder of this chapter is organized as follows. Section 1.1 discusses the motivation for empirically examining reviewer judgments and the likelihood of committing Type II errors. Section 1.2 summarizes the underlying theory used to motivate the study of several hypotheses regarding the impact of the external cues on the likelihood to commit Type II errors. Section 1.3 briefly overviews the research design and tests. Section 1.4 summarizes the organization of the remainder of the dissertation.

1.1 Motivation and Review of Working Paper Review Literature

The primary motivation of this study is to provide experimental evidence on whether the working paper review process can decrease the likelihood of committing

⁴ Respondents in Ameen and Strawser (1994) rated six analytical procedures on a seven-point Likert scale, with 1 indicating never used and 7 indicating always used. Simple analytical procedures and their ratings for Big Six firms were as follows: previous year’s balance, 6.4; scanning, 5.2; ratio analysis, 5.0; and trend analysis, 4.8. Complex analytical procedures and their ratings for Big Six firms were regression analysis, 2.0, and time-series analysis, 1.8. These results indicate a clear difference in usage between simple and complex analytical procedures. However, it is possible that complex analytical procedures are used for highly material items only.

Type II errors for substantive analytical procedures. Specifically, the study examines whether reviewers are more or less likely to commit Type II errors when given the requirement to provide written review notes and/or provided with a cognitive decision aid. The importance of Type II errors in this context is dependent on the significance of the working paper review process for substantive analytical procedures and the economic consequences of Type II errors. This section will discuss these areas and briefly note research findings to date.

Gillett, Bamber, Mock and Trotman (1994) argue that research on the review process has been limited to-date and has only begun to identify the nature of gains to the audit function. They also note that “there has been recognition of the importance of ... the review process since the early judgment studies (Gillett et al. 1994, p.35).” Furthermore, Rich, Solomon, and Trotman (1994) argue that current knowledge of reviewer behavior is at best, underdeveloped. Thus, they call for research that examines “reviewer cognition as well as the process by which task and contextual features result in purposeful alteration of the cognitive activities comprising audit review (Rich et al. 1994, p.1).”

Thus far, studies on the review process have focused on the ability of the review process to (1) increase the accuracy (identification of mechanical and conceptual errors) of information contained in the working papers, and (2) reduce the variance between judgments contained elsewhere within the working papers. Trotman (1985) and Trotman and Yetton (1985) found that the working paper review process was associated with decreased variability in control system judgments. Ismail and Trotman (1994) found that reviewers generated more alternative hypotheses than preparers. Ramsey (1994) found

that managers were better at detecting conceptual errors within working papers, while seniors were better at detecting mechanical errors. Also, Libby and Trotman (1993) found that reviewers were better than preparers at recalling information that was inconsistent with management explanations. Taken together, these studies suggest that the working paper review process can serve a valuable quality control function for the audit as a whole. The combined results also suggest that reviewers have more developed knowledge structures in place than preparers. These structures should better enable them to consider alternative explanations when reviewing substantive analytical procedures.

Recently, researchers have begun to examine the working paper review process from the standpoint of reviewer behavior. The studies mentioned above compared reviewer performance to preparer performance. However, there are unanswered questions that surround the determinants of reviewer performance. Rich, Solomon, and Trotman (1994) describe a cognitive model for the review process that sheds some light on possible determinants of performance. They describe the reviewer as the recipient of a stylized, persuasive message from both the client and preparer. Thus, Rich et al. (1994) argue that any study examining the working paper review process should consider the impacts of receiving persuasive messages and other environmental cues.

Rich (1996) tests the impact of perceived business risk and preparer competence on the behavior of the reviewer. Results from his experiment are not yet available; however, the experiment conducted in this dissertation controls for these environmental factors by holding them constant and measuring reviewers' perceptions of the factors to use in sensitivity analyses.

Asare and McDaniel (1995) examine the effects of task complexity and reviewer familiarity with the preparer within the review environment. They find that reviewers of working papers for which the reviewer is familiar with the reputation of the preparer detect more incorrect conclusions than reviewers of unfamiliar preparers in complex tasks only. This dissertation controls for Asare and McDaniel's (1995) results by making the preparer familiar to the reviewer because Type II errors should be more likely to occur in such circumstances.

The other branch of working paper review literature deals with the strategy undertaken by the reviewer when deciding how to review the working papers for an engagement and integrate the results. These studies have little to do with the review process for substantive analytical procedures alone. However, the results of these studies and this dissertation may be relevant for considering *when* to review analytical procedure results and *how* to consider the results in the context of other working papers examined during the audit. Thus, this literature also is briefly discussed.

Bamber, Bamber and Bylinski (1988) found that reviewers spend much of their cognitive processing time when conducting the review process on reading and understanding audit findings. They also found that analytical procedures performed during audit planning are an important source of information that reviewers use to form their perceptions of misstatement risk for various accounts. Bamber and Bylinski (1987) found that if reviewers become aware of high-risk accounts, they will be more likely to carefully scrutinize the audit evidence gathered for the account.

Phillips (1996) extended this research by identifying that reviewer performance was most effective when high-risk accounts were examined early in the review process. He defined reviewer performance as the detection of misstatements involving “gray-area” financial reporting issues. He argues that auditors are better able to integrate audit evidence when they review high-risk accounts early in the audit and then review lower-risk accounts (that may shed more light on misstatements once integrated) later in the audit.

Finally, Moeckel (1990) and Moeckel and Plumlee (1989) examined reviewer memory recall. Moeckel (1990) found that memory errors occur by reviewers when integrating audit evidence. More experienced auditors made reconstruction errors, which she defined as altering mental representations to make them consistent with existing knowledge. She also found that less experienced auditors failed to integrate more often, which she defined as a failure to make mental connections between separately received pieces of information. Moeckel and Plumlee (1989) found that auditors were more confident in their recognition of previously reviewed working papers when they accurately recognized evidence that they had seen and when they confused their own inferences with actually observed evidence. They concluded that auditors are often as confident in their incomplete and inaccurate recognitions as they are in their accurate recognitions (Moeckel and Plumlee 1989).

This dissertation examines a unique and more context-specific role of the working paper review process: its ability to reduce the risk of Type II errors for substantive analytical procedures. In this situation, the working paper review process is examined as

an efficient way to ensure that analytical procedures have been performed to the extent that alternative misstatement explanations have been ruled out by the preparer. The efficiency arises because the working paper review process serves other functions for the audit, such as ensuring that documented evidence is defensible (Rich, Solomon, and Trotman 1994) and as a means of training and communicating with preparers (Libby and Luft 1993).

Also, this dissertation is the first to test the impact on the review process of instruction to provide review notes. It is not clear based on prior research and professional guidance whether the review note process increases or decreases the risk of committing Type II errors. As a means of gaining a better understanding of the cognitive demands of the review note process, the interaction between instruction to provide review notes and the provision of a cognitive decision aid is examined. Because the decision aid is designed to ease demands on attentional capacity caused by the review process, the combination of both external cues can demonstrate the extent to which the task of writing review notes affects cognitive effort. Cognitive theory (described in more detail in the following subsection of Chapter One and in Chapter Two) suggests that any increase in the likelihood of committing Type II errors that results from the introduction of review notes to the use of the decision aid is due to an increase in the cognitive effort necessary to consider alternative misstatement explanations. Thus, this dissertation is able to add to existing literature on the working paper review process by (1) highlighting the cognitive difficulty of the review process (possibly made worse by review notes) and (2) demonstrating the difficulty that reviewers have considering misstatement explanations

when reviewing substantive analytical procedures. The study also demonstrates that the use of a cognitive decision aid that serves as a retrieval plan can help make this process less difficult, which decreases the risk of Type II errors.

There are several errors that auditors can commit when performing analytical procedures that could result in financial loss to the audit firm: Type I errors⁵ and Type II errors. The focus of this study is on the prevention of Type II errors, however, Type I errors are first discussed to demonstrate how the auditor considers the risks and impacts of such errors.

There are two controls built into the auditing process that minimize the risk of committing Type I errors. First, extensive follow-up procedures are performed to make sure that a misstatement has actually occurred before any information is communicated to users of the financial statements. These procedures are performed to prevent (1) loss of the client and (2) liability for lawsuits brought on by client or financial statement users if no misstatement actually occurred but auditors reported that one did occur. Second, the client is motivated to find and disclose evidence to disconfirm the hypothesized misstatement to prevent the auditor from disclosing it to financial statement users. Palmrose (1991, Appendix C) provides some evidence that these controls are effective. In her listing of cases brought to trial against auditors from 1960 - 1990, there were no

⁵ A Type I error occurs when the audit firm reports that a client made a misstatement when no such misstatement actually occurred.

cases that involved an auditing firm disclosing an error or irregularity that actually had not been committed by the client.

A Type II error, however, is problematic for auditors because clients may (1) be unaware that a misstatement occurred or (2) not disclose an intentional misstatement to the auditors. If users of the financial statements endure losses as a result of the error or irregularity, the auditors can be held partially liable to the users. In 31 of Palmrose's (1991) listing of court cases against public accounting firms that were based on alleged negligent auditing procedures, 16 verdicts in favor of the plaintiffs were reported at an estimated loss to public accounting firms of approximately \$79 million in total.⁶ The negligent auditing procedures of some, if not all, of these cases were possibly caused by the failure of the firms to find errors or irregularities not disclosed by the client (Type II errors), which a more effective review process may have helped to detect.

1.2 Overview of Hypotheses Examined

The working paper review process for analytical procedures can be described under Kennedy's (1993, 1995) debiasing framework as consisting of both effort (attention capacity and motivation) and data (internal and external). Because the review process is an explicit quality control process that can result in material economic losses to the firm

⁶ Amount based on best available information in Palmrose (1991, Appendix C). Accordingly, amounts are comprised of jury awards or reduced settlements, where applicable. However, some cases were subsequently settled for undisclosed amounts or were still under appeal. Thus, this estimate is conservative.

if not performed appropriately, an assumption is made that reviewers are motivated substantially beyond a zero level of effort to perform the working paper review.⁷ The unanswered question is whether they have sufficient capacity to perform the review, given the number of cognitive activities that reviewers perform when reviewing working papers. Furthermore, the role of the requirement to provide review notes is not known. The process of writing review notes may relieve the demands on reviewers by cueing them to counterexplain and consider alternative explanations. However, the review note process may be an additional cognitive task that increases the demands on the reviewer. The cognitive decision aid, however, was designed to relieve demands on capacity so reviewers could have more capacity available to consider alternative explanations. First, an hypothesis involving what role review notes play is discussed. Then, an hypothesis involving the decision aid is introduced. These hypotheses are followed by hypotheses that involve various comparisons and combination of the cues.

Because the instruction to provide review notes is currently used in practice, this dissertation attempts to understand better its effect on reviewers; whereas, the cognitive decision aid is designed to enable better data access for reviewers. The cognitive process of writing review notes should reduce the risk of committing Type II errors for reviewers who do not have excessive demands on capacity from performing the review. These

⁷ However, the experiment is designed to rule out motivation as a cause of any effects found between experimental conditions. Also, several measures for motivation are gathered and tested as covariates.

reviewers are able to write down possible working paper deficiencies, including counterexplanation of documented conclusions and the failure to rule out alternative hypotheses. According to the counterexplanation effect (Anderson et al. 1980, Koonce 1992), writing out arguments against the hypothesis should decrease the reviewer's likelihood of accepting the documented conclusions as complete. However, if there are excessive demands on a reviewer's capacity as a result of the review process, the instruction to explicitly state working paper deficiencies places even more demands on capacity. This increased strain on capacity can result in an increased likelihood of committing a Type II error because the reviewer focuses on other kinds of deficiencies and is less likely to consider alternative explanations.

Thus, hypothesis one involves first dividing subjects according to whether their review notes contain counterexplanation or alternative hypotheses not ruled out. Reviewers who provide review notes that contain either type of information are predicted to be less likely to commit Type II errors than reviewers who did not receive either external cue. Reviewers who do not provide any counterexplanation or alternative hypotheses in their review notes are predicted to be more likely to commit Type II errors than reviewers who did not receive either external cue.

The cognitive decision aid is designed to create a retrieval plan, which is an organized set of cues to aid the retrieval of information (Medin and Ross 1992). Because the decision aid provides a means for considering working papers in order to prevent Type II errors, the second hypothesis predicts that reviewers provided with the decision aid will be less likely to commit Type II errors than reviewers who only are instructed to

provide review notes. Support for this hypothesis would suggest that the decision aid works as intended: it enables the reviewer to internally organize the considerations for reviewing analytical procedures, which frees up capacity to consider alternative explanations. However, the possibility exists that the provision of a decision aid causes the reviewer to increase his or her motivation to consider alternative explanations.

Hypothesis five, discussed below, is designed to rule out this competing explanation.

Hypothesis three involves a comparison between the impact of provision of the decision aid to the instruction to write review notes. The hypothesis predicts that reviewers who only are provided the decision aid will be less likely to commit Type II errors than reviewers who only are instructed to write review notes. Should review notes cause an increase in likelihood to commit Type II errors, hypothesis three will hold necessarily because each cue will have opposite impacts on the likelihood to commit Type II errors. However, even if review notes lead to a decrease in the likelihood to commit Type II errors, the decision aid should have a stronger impact because it is designed to relieve demands on attentional capacity in addition to cueing consideration of alternative explanations.

Hypotheses four and five involve predictions for reviewers who are provided with both external cues. These hypotheses are needed to demonstrate the robustness of findings for hypotheses one and two. Also, these hypotheses help rule out competing explanations for results of the first two hypotheses.

Hypothesis four examines the impact of adding a decision aid to the current practice of instructing reviewers to provide review notes. The hypothesis predicts that the

provision of a decision aid will decrease the likelihood of committing a Type II error when compared to reviewers who only have been instructed to provide review notes. Although the prediction is the same in either case, the interpretation of the results changes depending on the impact of review notes found in hypothesis one. If review notes increase the likelihood of committing Type II errors, the provision of a decision aid should offset, or at least reduce, the negative effects of review notes. However, if review notes decrease the likelihood of committing Type II errors, the inclusion of a decision aid is predicted to relieve attentional capacity even further, which should increase the quantity and quality of review notes. In either case, the introduction of the decision aid is predicted to be beneficial. However, the former scenario suggests it eliminates a current problem in practice, and the latter scenario suggests it makes an effective process even more effective.

Hypothesis five is perhaps most important because it serves as a strong indicator of the role that review notes plays in the review process. The prediction for hypothesis five depends on the results of hypothesis one testing because it predicts the effect of adding review notes to the provision of a decision aid.

If review notes are found generally to aid data access, then Hypothesis 5 predicts that reviewers who receive both external cues should be less likely to commit Type II errors than reviewers only provided the decision aid. This prediction is based on the counterexplanation effect of writing the review notes. Because the reviewer has excess capacity to counterexplain and consider alternative hypotheses after reading the decision aid, they should be able to effectively counterexplain within the review notes. Thus, the

counterexplanation effect predicts that they will be less likely to commit Type II errors because they were able to clarify and write out reasons why the preparer's documented conclusions might not be true. This result, in conjunction with the results from hypothesis one, would provide strong evidence that instruction to provide review notes cues counterexplanation and that the review notes themselves proxy for counterexplanation.

If, however, review notes are found to increase the likelihood of committing Type II errors for hypothesis one, then Hypothesis 5 predicts that reviewers who receive both external cues are predicted to be more likely to commit Type II errors than reviewers who only are provided a decision aid. This prediction is based on the argument that the instruction to provide review notes represents the requirement of the reviewer to perform a difficult cognitive task in addition to counterexplanation. If the decision aid is found to decrease the likelihood to commit Type II errors in hypothesis two, then the additional requirement to provide review notes should not increase the likelihood to commit Type II errors, unless providing review notes is an additional cognitive task. In other words, the reviewer has sufficient capacity to provide written review notes because of the decision aid. Therefore, the requirement to provide review notes will only increase the likelihood of Type II errors if that requirement places an additional strain on capacity such that part, or all, of the benefit of the decision aid is eliminated. Thus, an increase in the likelihood to commit Type II errors for hypothesis five, in conjunction with results from hypothesis one, provides strong evidence that the instruction to provide written review notes is a difficult cognitive task. As a result, the possibility of Type II errors resulting from the

preparation and review of substantive analytical procedures would be a legitimate concern that auditing firms need to investigate further.

1.3 Overview of Research Design and Tests

This dissertation includes 78 audit managers and seniors from two national accounting firms in a between-subjects comparison of how two external cues affect the judgments of reviewers of analytical procedures working papers. The experiment included mainly reviewers with specialized knowledge in the financial services industry. This industry was chosen because one of the participating firms believed it to be an industry that relies heavily upon substantive analytical procedures. Necessary information to test whether subjects are likely to commit a Type II error is also gathered. Results of analysis show that those subjects who were more likely to commit Type II errors responded more strongly to the manipulations, which suggests that controlling for environmental factors is important for experiments involving the review process, as suggested by Rich, Solomon, and Trotman (1994).

The task involved providing all subjects with the same client and engagement background on a banking client. Subjects then received either one, both, or neither manipulation(s). Then, subjects reviewed a set of working papers involving the comparison of a key financial ratio for the client, which decreased from expectations materially. The working papers contained no errors or inconsistencies; however, they failed to rule out alternative misstatement explanations for the material fluctuation of the ratio. All subjects then made judgments regarding the overall quality of documented conclusions in the working papers. Finally, subjects listed alternative hypotheses;

responded to questions checking the manipulations and measuring their perceptions of the background information; and provided demographic information.

To test the hypotheses, the review notes provided by subjects had to be analyzed to determine whether they contained counterexplanation or alternative hypotheses. Two judges with public accounting experience who were not involved in the experiment classified review notes and found few that contained either type of explanation.⁸ Thus, all hypotheses were tested according to a model that suggests providing review notes may place added strain on attentional capacity. A 2 x 2 ANOVA model with contrast weights was used to test all five hypotheses. The results are consistent with the hypotheses, which suggest that (1) providing review notes places a strain on attentional capacity and (2) using a cognitive decision aid relieves the strain on attentional capacity and aids data access. However, when using both cues, the decision aid helps offset the additional strain on capacity encountered when providing review notes only.

1.4 Organization of the Dissertation

⁸ Furthermore, those few review note listings that contained any counterexplanation did not appear to be influenced by the counterexplanation evidenced by their dependent measure scores. Thus, it appears instruction to provide review notes does not cue counterexplanation. Also, the review notes are not a proxy for counterexplanation, and the predictions of the counterexplanation effect are not expected to hold. All hypotheses are therefore tested under the prediction that review notes are an additional cognitive task that increased the likelihood of committing Type II errors.

The remainder of the dissertation is organized into chapters. Chapter Two develops the underlying cognitive theory and summarizes previous empirical research to motivate five hypotheses about the impact of two external cues on reviewers' evaluations of analytical procedure working papers. Chapter Three describes the research design of the study, including the subjects. Chapter Four contains the analysis and results of experimental tests. Chapter Five includes a summary of the study and describes inherent limitations associated with the study's research design.

Chapter Two

THEORY AND HYPOTHESES DEVELOPMENT

This chapter develops the underlying theory for five hypotheses about the impact of two external cues, instruction to provide review notes and provision of a cognitive decision aid, on the likelihood of reviewers committing Type II errors. Such theory, which is based primarily on cognitive psychology theories of limited attentional capacity and information processing, suggests that unassisted review can be difficult because of the cognitive demands placed on reviewers of working papers.

Section 2.1 explicitly describes applicable cognitive theories and existing literature relevant for effective working paper review of substantive analytical procedures. Section 2.2 develops the hypotheses by discussing the theoretical basis for each of the external cues and the predicted impacts from introducing either one or both of the cues to the working paper review process. Section 2.3 describes important task and environmental considerations for the study. Section 2.4 summarizes the chapter.

2.1 Theories and Literature Relevant for Effective Working Paper Review

This section describes psychological theories and extant research findings that are important for better understanding of the working paper review process for analytical procedures. Section 2.1.1 describes theories related to cognitive effort, which Kennedy (1995, 1993) describes as consisting of capacity and motivation. Excessive demands on capacity, given fixed levels of motivation, is the foundation for the predictions in this dissertation. Section 2.1.2 discusses literature on alternative hypothesis generation and consideration. The findings from this literature are relevant because effective working

paper review for substantive analytical procedures includes ruling out alternative explanations that involve misstatements.

2.1.1 Cognitive Effort: Attention Capacity and Motivation

Kennedy (1993, 1995) developed a debiasing framework in an accounting context that models performance of tasks, such as analytical procedures, as a function of cognitive effort (see Figure 1 for an illustration of Kennedy's [1995] model).⁹ This dissertation assumes that effective performance of working paper review requires a significant amount of effort, given that the review of substantive analytical procedures is a complex task. Kennedy describes effort in terms of capacity and motivation. Thus, the following section will discuss the implications of capacity and motivation on the effectiveness of working paper review.

Kahneman's (1973) Capacity Model of Attention argues that there is a general limit on people's capacity to perform mental work or expend cognitive effort. His theory also stresses that individuals have a fair amount of control over how this capacity can be distributed across different task or task components. Kennedy (1993, p. 233) argues that

⁹ Kennedy's (1995, 1993) model also includes data quality as a function of performance. However, the predictions in this dissertation are not based on differences in the quality of data. Subjects in each experimental condition are exposed to the same external data, and subjects are randomly assigned to control for differences in internal data, which can be equated with knowledge. Knowledge differences between groups are then tested as a possible covariate.

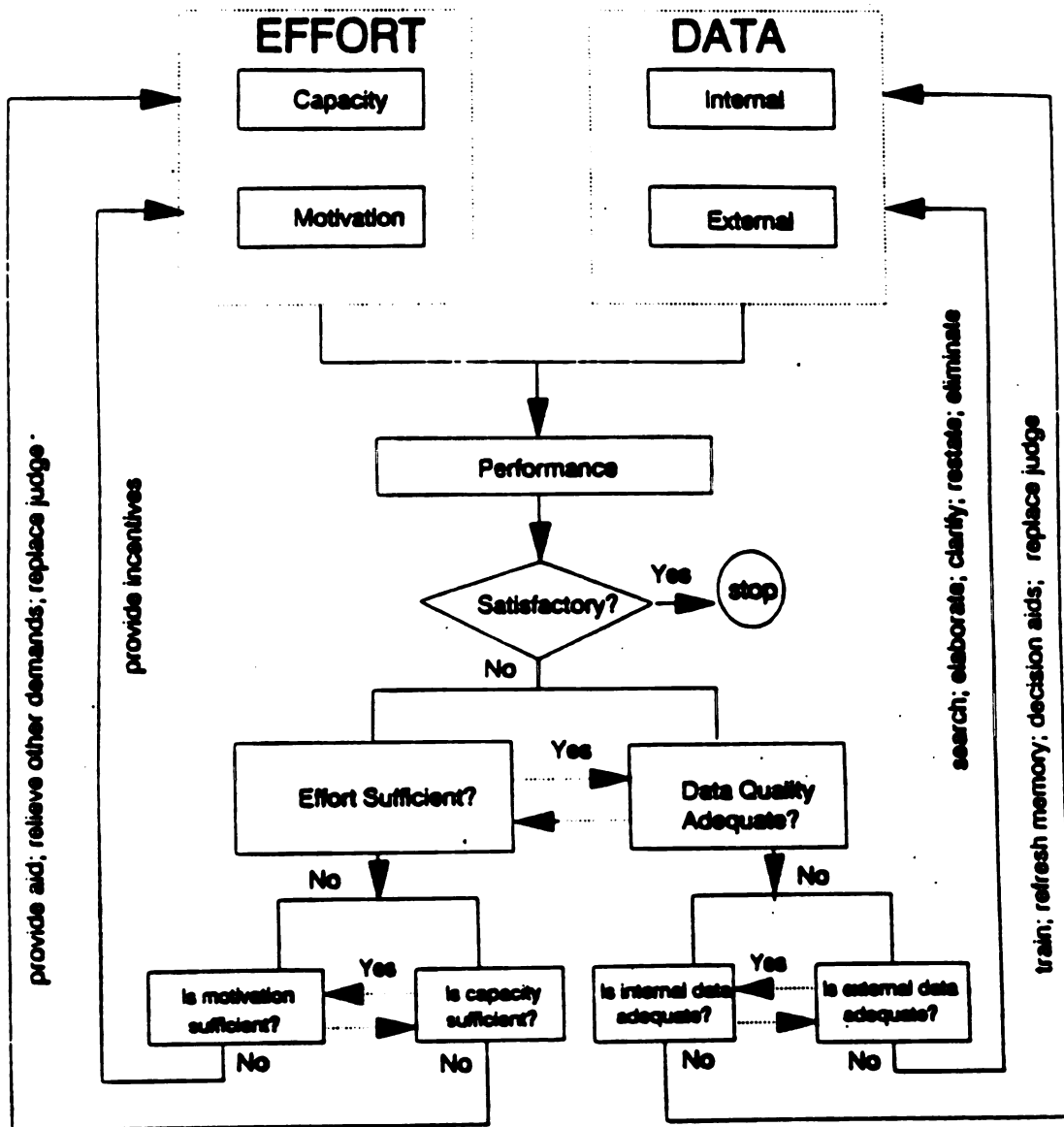


Figure 1
Kennedy's (1995) Debiasing Framework

“judgment performance may suffer when the task at hand demands more capacity than an individual has or can allocate to it.”

Shiffrin and Schneider’s (1977) model of Controlled Information Processing describes decision making as being an example of a controlled process. They argue that controlled processes require attention, which is bounded by limited capacity. Shiffrin and Schneider (1977, p. 160) further argue that “because these limitations prevent multiple control processes from occurring simultaneously, these processes often consist of the stringing together in time of a series of singly controlled unitary operations.” In simpler terms, decision-making requires individuals to allocate attention sequentially for complex tasks to adequately consider the individual components relating to the task.

Both Kahneman (1973) and Shiffrin and Schneider (1977) specify that attention is limited and can lead to a reduction in information search when available capacity is insufficient. Birnberg and Shields (1984, p. 367) note that these theories imply that “cognitive resource limitations of some sort are restricting the perceiver’s ability to attend to *all* stimuli.”

In an analytical procedures working paper review context, a reviewer must be able to perceive the account fluctuation, recognize the pattern involved, and retrieve from long-term memory the necessary information to evaluate appropriately the preparer’s hypothesis and any alternative misstatement hypotheses. Also, the reviewer must consider environmental factors surrounding the client, which include the likelihood of errors and fraud. Furthermore, the reviewer must consider the abilities of the preparer

and results from other areas of the audit. There exists the possibility that these tasks are difficult cognitively and, therefore, place a strain reviewers' attentional capacity.¹⁰

There has not been any explicit research in auditing concerning demands on attentional capacity. Birnberg and Shields (1984, p. 368) note that "it will require a reasonable amount of conjecture to generate the hypotheses about attention that are relevant to accountants." While this dissertation does not examine attention *per se*, it does focus on the reviewer's ability to attend to the considerations necessary for effective review of substantive analytical procedures.

While reviewers could have more knowledge than preparers of necessary analytical procedure considerations such as alternative hypotheses (c.f. Libby 1985, Libby and Frederick 1990, and Frederick 1991), retrieving and considering them adds another cognitive requirement to the review process. This requirement increases the demands placed on attentional capacity. It is an unanswered question, however, whether these demands on capacity place an excessive strain that reduces the ability to perform an information search of long-term memory or external data.

Kennedy's (1993, 1995) second component of cognitive effort, motivation, also can affect performance. An important issue is whether reviewers are motivated to

¹⁰ Also, there exists the possibility in a natural environment that the reviewer will not be focused totally on the review at hand. For example, the reviewer may be thinking about an issue at another client or a personal matter, either of which would reduce the attentional capacity available for the working paper review.

prevent Type II errors. Because working paper review is an explicit quality control mechanism designed to reduce the likelihood of Type II errors for substantive analytical procedures, it is assumed that reviewers are motivated to some extent to prevent Type II errors. Furthermore, there are costs in the form of litigation and reputation associated with failing to detect misstatements that are present in the financial statements (Palmrose 1991).

Rich (1996) argues that reviewer motivation is a function of the reviewer's perception of (1) the probability of undetected preparer error and (2) cost of judgments inconsistent with client or user desires. Rich defines undetected preparer error as the reviewer's belief of the likelihood that a misstatement exists, given that the preparer has not detected a misstatement. He defines the cost of judgments inconsistent with client or user desires to be the belief that detection of any misstatements may result in unfavorable conflict with the client. Such conflict can result in termination of the firm by the client (Gibbins and Emby [1985] describe this situation as client accountability). Rich (1996) argues that reviewer motivation will decrease as perceptions of undetected preparer error decreases and perceptions of cost of judgments inconsistent with client or user desires increase.

An important distinction between the reviewer's perception of undetected preparer error and the occurrence of a Type II error needs to be discussed. If a reviewer believes that there is not likely to be a misstatement in the financial statements, then the likelihood of an actual misstatement probably is lower than when the reviewer believes that a misstatement is likely. However, the existence of a misstatement is different than the

reviewer's perception that a misstatement exists. Thus, while the probability of a misstatement may be lower when the reviewer does not expect one, the likelihood that the reviewer will *not* detect one that does exist is greater because of the decreased level of motivation. Therefore, situations in which the reviewer perceives the likelihood of undetected preparer error as low actually increase the likelihood of Type II errors, given that a preparer error has occurred, because of decreased reviewer motivation.

Extant research in auditing has found that performance on analytical procedures improves when subjects are held accountable for their work. This accountability increases the subjects' motivation to perform (e.g. Lord 1990, Koonce et al. 1994). This dissertation controls for accountability in a manner similar to Koonce et al. (1994) by having reviewers put their names on their instruments. Reviewers were also informed that their work would be reviewed by Michigan State's accounting department and that there would be a discussion of the study within their experimental group.

However, making reviewers accountable does not help answer whether reviewers are motivated sufficiently to identify potential misstatements that preparers or clients have failed to detect. Unfortunately, limitations on subject availability prevent this study from manipulating situations in which reviewer motivation is expected to be different. Thus, the experiment in this dissertation was designed to elicit equal reviewer motivation across cells. However, the experiment was also designed so that given a misstatement, a Type II error is more likely to exist. Thus, the client is described such that the reviewer has a low perception of undetected preparer error. If reviewer motivation in this situation is not sufficient to perform the working paper review task, the manipulations used to test

the hypotheses developed in this chapter should not work as designed. Thus, if the hypotheses are supported, an argument can be made that reviewer motivation was sufficient.¹¹

2.1.2 Alternative Hypothesis Generation and Evaluation

Several accounting and psychology studies have documented subjects' difficulties in generating alternative hypotheses (see Bonner and Pennington 1991 and Koonce 1993 for reviews within the auditing context). Ruling out alternative explanations that involve misstatements is a critical component of the review process for analytical procedures. This phase is important because the evidence gathered by auditors for substantive analytical procedures is the primary evidence for the account(s) in question. When unexpected fluctuations occur, auditors need to obtain assurance that no misstatement occurred.

Libby (1985) and Libby and Frederick (1990) were the first studies in auditing to examine the importance of hypothesis generation and the impact of knowledge structure on the ability to generate alternative hypotheses. Studies documenting the frequency knowledge of auditors (c.f. Ashton 1991, Libby and Frederick 1990) found that auditors had difficulty generating error hypotheses beyond those errors that are expected to occur most often. Auditors can be exposed to plausible misstatements through training and discussions of misstatements from other auditors. However, the effectiveness of such

¹¹ The experiment also included measures of motivation that were used to rule out differences in motivation across cells.

exposure for enabling auditors to store and retrieve from memory plausible alternative explanations for account fluctuations is not known.

Koonce (1993) notes that failure to consider omitted hypotheses may be due to the inability to retrieve the hypotheses, as suggested by Mehle (1982) and others (Fischhoff et al. 1978 and Mehle et al. 1981, for example). In an auditing task involving preparers of analytical procedures, Bedard and Biggs (1991b) found that preparers appeared to be effective at perceiving and recognizing patterns in account fluctuations; however, their results also suggested that auditors have difficulty in retrieving appropriate hypotheses. Heiman (1990) found that many of the preparers used in her experiment examining the evaluation of client-prepared explanations for analytical procedure fluctuations failed to generate more than two alternative hypotheses. Those subjects then increased their likelihood judgments for a client's hypothesized cause. A similar result also was found in a psychology study by Mehle et al. (1981).

In a similar auditing study, Bedard and Biggs (1991a) found that 12 of their 42 preparer-subjects adopted incomplete client-provided hypotheses, most likely because of the failure to adequately consider alternative hypotheses. As noted in Bedard and Biggs (1991a), this difficulty was reduced but not eliminated by more "expert" auditors.

A similar finding to Bedard and Biggs (1991a) also was documented in a cognitive psychology study by Mehle (1982). He found that subjects could generate only a partial list of alternative hypotheses; however, they believed that their lists were complete. Although Mehle (1982, p. 104) notes that it is not necessarily optimal to carry an exhaustive set of hypotheses through the decision-making process, he argues that "in

applied settings, there exists a large class of decision problems which require the decision maker to generate exhaustive, or nearly exhaustive hypothesis sets.”

Although a reviewer need not consider all hypotheses for an account fluctuation, it can be argued that when presented with a nonmisstatement explanation, a reviewer should be able to consider and rule out at least the most frequent hypotheses that involve a material error or fraud on the behalf of the client. Bedard and Biggs (1991a, p. 88) noted, “[Auditors’] failure to disconfirm is troubling if such behavior were to be found in a situation involving fraud. Thus, further research is needed to better understand how auditors evaluate management representations.”

The question that is relevant for this study is whether reviewers retrieve and evaluate alternative misstatement hypotheses as part of the working paper review process. Ismail and Trotman (1994) found that reviewers generated more alternative hypotheses than preparers, which suggests that reviewers have better developed knowledge structures of types of misstatements. However, it is not known whether reviewers are able to retrieve and consider such misstatements when performing working paper reviews of analytical procedures, given the possible demands on attentional capacity caused by the review process.

The main implication of not ruling out alternative explanations that involve misstatements for the working paper review process is the increase in the likelihood of committing Type II errors. The working paper review process is the primary quality control mechanism for substantive analytical procedures. If a misstatement exists in the account(s) in question and neither the client or preparer has identified the misstatement in

the working papers, then the reviewer represents the “last line of defense” against allowing the misstatement to go undetected. Therefore, it is critical that the reviewer be able to effectively rule out alternative explanations that involve misstatements.

2.2 Development of Hypotheses

This section develops hypotheses that indirectly examine the concepts discussed in the previous section by directly testing the impacts of two external cues on the likelihood of reviewers to commit Type II errors. These cues, instruction to provide review notes and provision of a cognitive decision aid, can be used to help reviewers access stored knowledge by providing a means to organize existing knowledge on how to review substantive analytical procedures effectively. Furthermore, the design of the cues may shed light on the question of whether reviewers are experiencing a strain on available attentional capacity when reviewing working papers and whether the cues are able to relieve or increase this strain, if it appears to exist.

2.2.1 The Impact of Review Notes

One way to examine whether the review process is affected by excessive demands on capacity is the procedure of writing review notes. Because the use of analytical procedures involves hypothesis testing based on simple computations (Ameen and Strawser 1994), it is rare that an auditor will be able to support a given explanation. Rather, a more effective method for confirming a hypothesis is by ruling out competing explanations that involve an error or irregularity.

One difficulty for reviewers analyzing substantive analytical procedure conclusions is the number of other considerations that reviewers must address as part of

the working paper review. Reviewers must review the working papers to ensure that they are legible, coherent, and mechanically accurate. Also, reviewers analyze the working papers to determine whether the information contained within the papers is internally consistent. They are also concerned with how the working papers in question fit into the cycle being tested and the audit as a whole. Furthermore, reviewers must address the extent to which they believe the client or the preparer is knowledgeable about the information contained in the working papers and motivated not to report truthfully. Thus, a reviewer might have difficulty considering alternative misstatement hypotheses without being cued because of these demands on their attentional capacity.

The use of instructions to provide review notes when reviewing analytical procedures could cue the reviewer to consider alternative explanations, which might better enable the reviewer to retrieve and evaluate stored data. One way that writing review notes could serve this function is through the counterexplanation effect (Anderson et al. 1980, Koonce 1992). This effect predicts, according to Koonce (1992), that the increased causal reasoning accompanying counterexplanation development typically causes an event that is part of a written counterexplanation to be judged less likely to have occurred than an event that is not part of a written counterexplanation. Koonce found that counterexplanations supporting why client-provided representations might not be true decreased likelihood judgments.

The counterexplanation effect has two important implications. First, individuals seldom spontaneously counterexplain once a written explanation has been created. In a psychology study, Koriatic et al. (1980) found that individuals spontaneously generated

supporting reasons for their position but had to be explicitly instructed to provide information against their position. Second, when individuals are specifically asked to write a counterexplanation, information against the hypothesis in question becomes more developed and salient in memory. Thus if reviewers counterexplain when writing review notes, then retrieval of internal knowledge of possible alternative explanations could be enhanced.

Koonce (1992) found support for both implications of the counterexplanation effect for preparers of analytical procedures. Her preparers had difficulty counterexplaining evidence before first providing a written explanation for why the hypothesis provided by a client might be true. This result suggests that preparers have difficulty spontaneously counterexplaining. Koonce (1992) found, however, that counterexplanation supporting why client-provided representations might not be true decreased preparers' likelihood judgments. These findings suggest that instructions to provide review notes might be effective for helping to prevent Type II errors if reviewers can overcome the difficulty of developing counterexplanations.

Whether instruction to provide review notes proxies for explicit instruction to counterexplain is not known. This question is important because the current review process for analytical procedures includes writing review notes. For example, when working papers include a nonmisstatement explanation by the client and the preparer does not provide enough evidence to support management's explanation such that other misstatement explanations are ruled out, will the reviewer spontaneously counterexplain? If written review notes for such working papers contain explicit counterexplanation or

alternative hypotheses not considered and ruled out by the preparer, it is likely that review notes do proxy for explicit instruction to counterexplain. Thus, review notes may be a proxy for counterexplanation and the current review process might help reduce the risk of Type II errors.

However, if these review notes do not contain counterexplanation or alternative hypotheses not ruled out, then the instruction to provide review notes likely does not proxy for explicit instruction to counterexplain. Under this condition, the current review process may not help reduce the risk of Type II errors.

If instruction to provide review notes does not cue reviewers to counterexplain, then the instruction and subsequent review note preparation must have some impact on the reviewer. Because reviewers are not being benefited by the instruction to provide to review notes, there is a high likelihood that these instructions are perceived by reviewers to be an additional task to perform. Thus, the demands on attentional capacity associated with the review process are increased because reviewers must (a) focus on the *specific* aspects of the working papers that are not satisfactory, for whatever purpose, and (b) document them as review notes. Also, these review notes must be detailed enough to direct the preparer on how to address the deficiencies.

Thus, another possible outcome of instruction to provide review notes could impact the process negatively: the increase of demands on attentional capacity caused by the requirement to consider explicitly any working-paper deficiencies. Such consideration needs to occur to the extent that the reviewer can write out his or her perception of the deficiency and any follow-up procedures that need to be performed.

Prevailing theories of attentional capacity (c.f. Kahneman 1973, Shiffren and Schneider 1977) specify that attention is limited and can lead to a reduction in information search when available capacity is insufficient. The process of evaluating every aspect of the working papers and determining what needs to be done to improve documentation or better support the current explanation might require much or all of the reviewer's available capacity.¹² This strain of capacity could inhibit the reviewer from considering alternative hypotheses involving errors and irregularities not ruled out by the evidence gathered in the working papers.

Should instruction to provide review notes increase capacity demands, the risk of Type II errors will increase for reviewers of substantive analytical procedures. Reviewers must consider and rule out alternative misstatement hypotheses for Type II errors to be prevented. The likelihood of these hypotheses being considered will decrease if available attention capacity is insufficient. Thus, the greater the demands on capacity, the less likely reviewers will be able to consider alternative hypotheses.

¹² In addition, reviewers in a natural environment also might be considering issues related to other clients or professional matters. Furthermore, personal issues might be diverting the focus of a reviewer. Either of these distractions decrease available capacity needed to consider alternative explanations. Because these considerations are not expected to differ across experimental conditions and would strengthen the impact of the manipulations, they are not addressed in this dissertation.

Instruction to provide review notes can be placed into Kennedy's (1995) debiasing framework under either condition: it can cue counterexplanation or increase strain on capacity. However, the cue's impact on performance is constrained by the available capacity of the reviewer. If instruction to provide review notes proxies for explicit instruction to counterexplain and the reviewer has sufficient available capacity, then the process of writing review notes should reduce the risk of committing Type II errors. In other words, when a nonmisstatement explanation is documented in the working papers, a reviewer should make less favorable judgments of those conclusions if alternative misstatement explanations have not been ruled out. If the reviewer has sufficient capacity and the instruction to provide review notes does not proxy for counterexplanation, then there still should be no impact from writing review notes because the reviewer has sufficient capacity to consider counterexplanation. However, should instruction to provide review notes cause reviewers to make more favorable judgments of the documented conclusions in the same working papers, then it is likely that the cue results in excessive demands on reviewers' attentional capacity.

One way to test these competing possibilities is to compare reviewers who are asked to provide and who subsequently write review notes to those who are not provided with such instructions. If reviewers include counterexplanation or alternative explanations, then the instruction to provide review notes might serve as a cue to counterexplain. Thus, reviewers should be less likely to commit Type II errors. If, however, reviewers do not include counterexplanation or alternative explanations, then the instruction to provide review notes does not appear to serve as a cue to

counterexplain. Thus, the review note process is an additional cognitive task that should place an incremental strain on attentional capacity, and reviewers should be more likely to commit Type II errors. This comparison leads to the following hypotheses:

*H1a: Reviewers who are instructed to provide review notes and who subsequently write notes that include counterexplanation or alternative hypotheses are **less** likely to commit Type II errors than reviewers who have not been given such instructions.*

*H1b: Reviewers who are instructed to provide review notes and who subsequently write notes that do **not** include counterexplanation or alternative hypotheses are **more** likely to commit Type II errors than reviewers who have not been given such instructions.*

2.2.2 A Retrieval Plan to Relieve Demands on Capacity

If the review process is cognitively challenging, an important question to be addressed is whether the relief of demands on capacity will better enable reviewers to consider the adequacy of documented conclusions. Shiffren and Schneider (1977, p. 156) argue that “controlled processing of [memory] node activation (i.e., access to long-term memories) requires active attention. Controlled processes are therefore tightly capacity-limited.” Shiffren and Schneider go on to argue that automatic processing can overcome these difficulties. However, substantive analytical procedures are unstructured and unique to each client. Thus, automation, which occurs as a result of repetition, is not likely to occur for this auditing task. Therefore, the review process is predicted to require the type of controlled information processing that Shiffren and Schneider argue is limited by available capacity.

One important property of Shiffrin and Schneider's (1977) theory of controlled information processing is that these controlled processes can be adopted quickly, without extensive training, and modified fairly easily. Thus, reviewers should be able to respond to mechanisms designed to ease demands on capacity by better structuring the information to be processed by controlled processes.

Ashcraft (1994, p.238) notes that "access is increased by effective retrieval cues." Medin and Ross (1992, p.186) suggest that recall of relevant information is much higher if an individual has a retrieval plan, which they define as "an organized set of cues to use in helping to retrieve information." They further stress that such a plan for retrieval should include two main characteristics. First, the various cues should be organized in such a manner that individual cues are not forgotten. Second, the individual cues in the plan should be able to lead to the recall of items not recallable from other cues and such recall should occur in a systematic fashion. For instance, when asked to name the 50 states, most individuals break the country into geographic regions and then name states in the region (Medin and Ross 1992). Using this strategy, the most difficult states to name are those that lie on the border of regions (e.g., is Montana a Northwest or Great Plains state?).

Furthermore, Shiffrin and Schneider's (1977) controlled processing model suggests that systematic processing relieves strains placed on an individual's fixed level of capacity. In fact, their model argues that limitations of controlled processing prevent multiple control processes from occurring simultaneously. Instead, these processes consist of stringing together in time a series of singly-controlled unitary (i.e., singular)

operations. Thus, the use of a retrieval plan can help relieve demands on capacity by providing a systematic process for evaluating analytical procedures.

An auditor reviewing analytical procedures should benefit from having a cognitive retrieval plan. A reminder cue that contains a brief summary of the general procedure for reviewing analytical procedures can benefit the reviewer in several ways. According to Kennedy's (1993, 1995) debiasing framework, this type of decision aid can help a reviewer ease capacity strains by providing the reviewer with a systematic method of reviewing analytical procedures. Extant auditing studies have shown that decision aids improve auditor performance (Bonner, Libby, and Nelson 1996 and McDaniel and Kinney 1995). An unanswered question is whether a decision aid affects reviewers of analytical procedures who have received an inherited hypothesis from the preparer.

The use of an informal, unstructured aid is appropriate given the subjective nature of the review process and, perhaps most importantly, analytical procedures. A checklist format is not used because it can lead to output interference caused by the reviewer anchoring the review only on the checklist (Pincus 1989). McDaniel and Kinney (1995) argue that informal, unstructured aids are appropriate for analytical procedures because (1) auditors use subjective, judgmental methods to apply analytical procedures and (2) there are perceived benefits from matching the decision aid structure to that of the task. Also, the decision aid should not be a list of specific possible alternative hypotheses. As discussed in Heiman (1990, p. 88), "presentation of a partial list of alternatives could inhibit retrieval of additional alternatives (i.e., it could cause output interference)."

The decision aid used for reviewers of substantive analytical procedures should include a sequential list of independent reminders that mirror the judgments reviewers normally make. For this dissertation, the decision aid was developed through consultation with firm experts and tested using firm experts and reviewers during a pilot test.¹³ Thus, the decision aid included the following consideration reminders: the motivations of the client and preparer, alternative explanations, comparisons of current-year data to prior years and industry, all components of any analytical procedures based on ratios or models, relationship to transaction cycles and management assertions, causes of the fluctuations (general, not specific), and procedures for follow-up of unexpected results. The exact decision aid used is shown in Figure 2.

Therefore, the guidance should cue the reviewer to recall memories relating to the general types of reasons for account fluctuations (e.g., trends or seasonality), both misstatement and nonmisstatement. When a nonmisstatement explanation is the hypothesis inherited from the preparer, the reviewer can then decide whether (1) alternative misstatement explanations have been considered appropriately and ruled out or (2) the preparer has not explained completely or appropriately the fluctuation based on the information provided.

¹³ For further information concerning the design and testing of the decision aid, see Chapter 3, Research Design.

Please read and consider the following while performing the review:

IMPORTANT REMINDERS TO CONSIDER WHEN REVIEWING ANALYTICAL PROCEDURE RESULTS

When reviewing analytical procedure results that contain any significant fluctuations from auditor expectations, it is important to consider two risks. First, the explanation given by the *client* to the preparer of the working papers may not appropriately explain the fluctuation. The explanation may be valid, but it may not account for the material portion of the fluctuation. Also, there is a chance that the explanation may not be valid. Second, the explanation provided by the preparer of the working papers may not be corroborated to the extent necessary to explain the material portion of the fluctuation (or it may not properly reflect the explanation provided by the client).

Therefore, it is important as the reviewer to consider alternative explanations to the cause(s) of the fluctuation provided by the client to the preparer of the working papers. Keep in mind that fluctuations are often caused by trends, seasonal patterns, changes in the economy, movements in other financial data, errors, and, in some cases, irregularities. When considering such alternatives, start by examining what you know about the client both in the current and past fiscal years, including important comparisons with other companies in the same industry. Then, consider the components of the analytical procedure and their relationship to other accounts or ratios. Remember to consider key aspects of the related internal control structure and the management assertions covered by the analytical procedure.

Proper procedures for following up significant fluctuations from auditor expectations include discussion with management, evaluation of the reasonableness of such discussions, and corroboration of client explanations for the fluctuations. The evidence obtained should directly support the explanation and be quantified such that the material portion of the fluctuation is explained.

Figure 2
Decision Aid Designed to Relieve Demands on Capacity

The use of the decision aid is predicted to relieve demands on attention capacity by enabling the reviewers to use a retrieval plan to more efficiently process information. Thus, reviewers are predicted to have more available capacity to consider the extent to which the evidence supports the documented conclusions in the working papers. As a result, reviewers should be less likely to accept a nonmisstatement explanation that is not sufficient enough to rule out competing misstatement explanations. Therefore, reviewers should be less likely to commit Type II errors when first reading the decision aid. A comparison of reviewers provided with a retrieval plan with those who are not provided with such a plan leads to the following hypothesis:

H2: Reviewers who are provided with a decision aid are less likely to commit Type II errors than reviewers who have not been provided with such an aid.

Because the decision aid provides reviewers with an external data source (guidance for reviewing analytical procedures) that relieves demands on capacity while instructions to provide review notes might increase demands on capacity, using a decision aid is predicted to have a stronger impact than writing review notes. Even if instruction to provide review notes cues counterexplanation, the decision aid is predicted to have a stronger effect because of the increased available capacity to consider alternative explanations. No such prediction is made regarding the instruction to provide review notes. Thus, these arguments lead to the following hypothesis:

H3: Reviewers who are provided with a decision aid are less likely to commit Type II errors than reviewers who are instructed to provide review notes.

2.2.3 The Combination of Review Notes and a Decision Aid

Comparisons of groups with both manipulations to those groups with only one manipulation can be performed. Such tests can help better examine the effect of review notes on the review process and whether demands on capacity are reduced when the decision aid is used. If the decision aid relieves demands on capacity, then reviewers with both manipulations should be less likely to commit Type II errors than reviewers who have only the review note manipulation. Likewise, the effect of review notes on the review process can be more clearly observed by comparing reviewers with both manipulations to reviewers who only were provided the decision aid. The following discussion and hypotheses will develop these concepts more thoroughly.

To test the robustness of the decision aid's impact on the review process, reviewers that only are provided with instruction to provide review notes can be compared to reviewers provided with both a decision aid and instruction to provide review notes. Because the decision aid was designed to relieve demands on capacity, this comparison should result in more available capacity for the combined condition group who were provided with the decision aid. Thus, reviewers provided with both manipulations are predicted to be more likely to consider alternative explanations and the sufficiency of documented conclusions. As a result, reviewers in the combined manipulation condition should be less likely to commit Type II errors, which leads to the following hypothesis:

H4: Reviewers who are provided a decision aid and are instructed to provide review notes prior to evaluating an analytical procedure are less likely to commit Type II errors than reviewers who are instructed to provide review notes only.

Perhaps the most important comparison for identifying the effect of instruction to provide review notes on the review process involves reviewers provided with both manipulations and reviewers only provided a decision aid. If instruction to provide review notes is not an additional cognitive task that requires effort, then, given a decision aid to relieve demands on capacity, there should be no difference between the groups. However, if instruction to provide review notes cues counterexplanation, then reviewers who are provided with both manipulations should be less likely to commit Type II errors. This result is predicted because the group provided only with the decision aid has not been provided with the counterexplanation-inducing review note cue.

Also, if the decision aid creates a reduction in the likelihood of Type II errors because of an experimental demand effect, reviewers provided with both manipulations should be no different than reviewers provided with the decision aid only.¹⁴ If anything, the demand effect should be stronger for reviewers with both manipulations. This argument is made because such reviewers are also instructed to provide review notes,

¹⁴ The decision aid could induce an experimental demand effect if reviewers were more skeptical of the working papers reviewed simply because of the presence of a “warning” in the experiment, not the content of the aid itself.

which is current practice. Thus, the aid should be more likely to be interpreted as a “warning” in this condition.

However, should reviewers exposed to both manipulations be more likely to commit Type II errors than reviewers only provided a decision aid, there is only one explanation that makes sense. If the review process is cognitively challenging and requires a significant level of attentional capacity, then the additional instruction to provide review notes can be an additional, cognitively challenging task. Thus, even though the decision aid provides a set of systematic retrieval cues, the instruction to provide review notes may offset the benefits of the decision aid to some extent. The degree of elimination of the decision aid’s effect is an indication of the demands on capacity placed by the instruction to provide and subsequent writing of review notes.

The last hypothesis depends to some degree on the results of testing the first hypothesis regarding the impact of review notes. If instructions to provide written review notes are found to cue counterexplanation, then the combined effect of using a decision aid and providing review notes should be stronger incrementally than only using a decision aid. This prediction is based on the argument that the decision aid should reduce capacity demands and enable more consideration of alternative explanations. However, if providing review notes does not proxy counterexplanation, then the use of a decision aid alone is predicted to have a stronger effect than the combined effect of using both external cues. This result is expected because the instruction to provide review notes is an additional cognitive task, which increases demands on capacity. These conflicting predictions lead to the following hypotheses:

If instruction to provide review notes generally found to cue counterexplanation in

Hypothesis 1 testing:

H5a: Reviewers who are (1) provided a decision aid and (2) instructed to provide review notes are less likely to commit Type II errors than reviewers who have been provided a decision aid only.

If instruction to provide review notes generally found not to cue counterexplanation in

Hypothesis 1 testing:

H5b: Reviewers who are (1) provided a decision aid and (2) instructed to provide review notes are more likely to commit Type II errors than reviewers who have been provided a decision aid only.

When all five hypotheses are combined, there are two different models that are established based on the classification of review notes. If review notes are found to contain counterexplanation or alternative explanations, then instruction to provide review notes is argued to cue counterexplanation. Thus, a model that includes hypotheses 1a, 2, 3, 4, and 5a is predicted to hold. If, however, review notes are not found to cue counterexplanation or the consideration of alternative explanations, then instruction to provide review notes is posited to be an additional cognitive task that increases the likelihood of committing Type II errors. Therefore, a model that includes hypotheses 1b, 2, 3, 4, and 5b is predicted to hold. A summary of the hypotheses under both models is provided in Figure 3.

	No Instructions to List Review Notes	Instructions to List Review Notes
No Decision Aid	Cell A	Cell B
Decision Aid	Cell C	Cell D

Model A	Model B
<u>Review Notes Aid Data Access</u>	<u>Review Notes Strain Capacity</u>
H ₁ : Cell B < Cell A	H ₁ : Cell B > Cell A
H ₂ : Cell C < Cell A	H ₂ : Cell C < Cell A
H ₃ : Cell C < Cell B	H ₃ : Cell C < Cell B
H ₄ : Cell D < Cell B	H ₄ : Cell D < Cell B
H ₅ : Cell D < Cell C	H ₅ : Cell D > Cell C

Dependent Measure: Likelihood of Committing a Type II error

Figure 3
Summary of Hypotheses

2.3 Task and Environmental Considerations

The following section describes important environmental and task-specific considerations that must be considered for this dissertation. Specifically, belief adjustment for inherited hypotheses, *a priori* likelihood of committing Type II errors, and the processing strategy used by reviewers are discussed to the extent they need to be addressed in the design of the study.

2.3.1 Belief Adjustment for Inherited Hypotheses

Because a reviewer first processes the preparer's conclusion (before developing his or her own), it becomes "inherited" as an initial hypothesis by the reviewer, provided the reviewer has confidence in the preparer's integrity and ability (see Section 2.3.2). One reason for such an adoption is that the reviewer knows that the preparer had direct contact with the client and based the documented conclusion on such discussions. Thus, while the reviewer has an objective of ensuring that the working papers have been prepared appropriately, the stated hypothesis of the preparer becomes the point of departure for the reviewer.

The process of searching for internal data to evaluate an inherited hypothesis can be affected by the order in which reviewers evaluate evidence or the knowledge schema activated by the reviewer. Hogarth and Einhorn's (1992) belief adjustment model predicts that reviewers who make an evaluation of evidence after first reviewing all of the pieces of information in a complex task place too much weight on early evidence evaluated. This processing of early evidence is caused by the use of step-by-step processing (Hogarth and Einhorn 1992), which is an internal form of anchoring and

adjustment. Individuals compare each piece of evidence with the existing belief.

Inheriting a nonmisstatement hypothesis from the preparer could cause a reviewer to be more likely to adopt the hypothesis without considering alternative hypotheses when the evidence items reviewed in the working papers do not contradict the conclusion reached in the working papers by the reviewer. This processing can result in a primacy effect, which would increase the likelihood of a Type II error when a nonmisstatement conclusion is inherited by the reviewer.

The process of searching for internal data to evaluate inherited hypotheses can also be affected by the schema activated by the reviewer when examining working papers. Libby and Trotman's (1993) results can imply that reviewers have developed different knowledge structures (or schemas) from preparers, which affects their evaluations. They found that reviewers recalled more information inconsistent with a client's hypothesized cause after studying the information used by preparers when evaluating the client's hypothesized cause. Thus, the schemas activated by an inherited hypothesis may affect subsequent evaluation of arguments in a manner not considered by Hogarth and Einhorn (1992).

An important question not addressed in Libby and Trotman (1993) are the impacts of inheriting an hypothesis for which inconsistent information with the activated schema is not present. Thus, it is not known whether a reviewer who does not find inconsistencies with an activated schema would then attempt to counterexplain or generate alternative hypotheses not ruled out by the information presented.

Other possible problems associated with the evaluation of a potentially-biased, inherited hypothesis are the curse of knowledge and output interference. McDaniel and Kinney (1995) found that a preparer's effectiveness in performing analytical procedures is improved when given only instructions to form and document expectations and no book values. Kennedy (1995) also documented that auditors in an analytical procedures task were unable to ignore outcome knowledge when estimating what sales levels other auditors would predict. Furthermore, Heiman (1990) and Libby and Frederick (1990) suggest that providing preparers with an hypothesis might limit their ability to consider other hypotheses due to the robust psychological effect of output interference. Because a reviewer inherits from a preparer an hypothesis that contains book values, outcome information, and the preparer's hypothesis, there is some question whether, given the appropriate motivation, the reviewer will be able to effectively perform the review by considering and ruling out alternative explanations.

2.3.2 *A Priori* Likelihood of Committing Type II Errors

Rich, Solomon, and Trotman (1994) describe the reviewer as the recipient of a persuasive message. Reviewers must process the client's persuasive message in the form of references to client explanations in the preparer's working papers. Professional skepticism requires that the auditor must be cognizant that a client explanation may not be accurate.¹⁵ Two biases that a reviewer should be concerned with are (1) reporting bias,

¹⁵ Of course, professional skepticism does not imply that a management explanation is to be assumed to be misleading. However, it is reasonable to assume that

the degree to which a communicator's willingness to convey an accurate message is compromised, and (2) knowledge bias, the degree to which a communicator's knowledge of message relevant information is inaccurate (Eagly and Chaiken 1993, 1984).

For the review process, reporting bias can be exhibited by either the client or the preparer. Clients may not convey an accurate message to an auditor when an irregularity (i.e., fraud) has occurred. Also, if a manager or employee of the client realizes an error has occurred, the individual may not want to be exposed by the error. Thus, misrepresentation can occur for any misstatement committed by the client. Preparers also can exhibit strategic behavior that includes documenting inaccurate messages in the working papers (Rich 1996, Rich et al. 1994). These misrepresentations can occur for various reasons. However, the most likely misrepresentation from a staff is "tick mark fraud," which occurs when a staff documents in the working papers that an audit procedure was performed when the preparer actually did not perform the procedure. Environmental factors, such as time pressure, or intrinsic factors, such as work aversion, are likely determinants of tick mark fraud.

In the task performed by reviewers in this dissertation, there is no indication from the environmental factors presented that the preparer or client has any motivation to provide inaccurate information. However, professional skepticism towards the client still

remaining neutral in opinion prior to accumulating evidence means that an auditor must allow for the possibility that a client's explanation is inaccurate.

is relevant. Thus, the study controls for factors that may be associated with reporting bias (See Chapter 3, Research Design, for further discussion of specific controls).

For the review process, knowledge bias also can be exhibited by either the client or the preparer. As previously mentioned, knowledge bias is the extent to which the knowledge of the communicator is inaccurate. Thus, reviewers are concerned with the knowledge of both the client and the preparer relating to the underlying information contained in the working papers. For explanations or unexpected analytical procedure fluctuations, the reviewer is concerned that the client has sufficient knowledge of the accounts, ratios, or relationships in question. In other words, the reviewer is concerned with whether the specific individual or individuals who provided the preparer with the explanation had direct knowledge or was “guessing”. Also, given that a relatively inexperienced staff may be performing the substantive analytical procedure, the reviewer is concerned with whether the preparer asked the right questions, understood the client responses sufficiently, and gathered sufficient, corroborating evidence.

In the task performed by reviewers in this dissertation, there is no indication from the environmental factors presented that the preparer or client lacked sufficient knowledge to provide or interpret the explanation given for the fluctuation. Also, the study controls for factors that may be associated with knowledge bias (See Chapter 3, Research Design, for further discussion of specific controls).

When reporting and knowledge bias is assessed as low, the risk of Type II errors is expected to be higher. This prediction is made because when nonmisstatement explanations are provided by clients and preparers, reviewers are expected to be less

likely to consider other explanations for the unexpected fluctuation. In other words, the reviewer has no *a priori* reason to suspect that the documented conclusions are inaccurate. Therefore, when an incorrect nonmisstatement explanation is documented by the preparer, the reviewer is less likely to detect the mistake. This scenario results in a Type II error.

When the reviewer assesses reporting and knowledge bias as high, any documented conclusion--especially nonmisstatement explanations--will be evaluated with a higher level of skepticism by the reviewer. When the reviewer believes that either the client or preparer has incentive to lie or is not knowledgeable, there will be more effort by the reviewer to consider the sufficiency of the evidence provided and alternative misstatement explanations. Because of the increased skepticism, these considerations are expected to dominate other working paper review considerations. Thus, excessive demands on attentional capacity should not lead to increased Type II error risk; rather, other areas of the audit (documentation, cohesiveness of working papers, etc.) are expected to receive less attention.

Because of limitations on subject availability, however, this dissertation only focuses on situations where Type II errors are likely to occur. This choice was made because of the economic consequences to firms for failing to detect misstatements through evidence gathered subsequent to an unexpected, analytical procedure fluctuation. Thus, the experiment was designed to elicit low levels of reporting bias--no expectation of lying--and knowledge bias--no expectation of insufficient knowledge--for the client

and preparer. While reviewer behavior across reporting and knowledge bias levels is an interesting topic, this question will have to be addressed in a subsequent study.

2.3.3 Reviewers' Information Processing Strategy

While effective working paper review of substantive analytical procedures includes ruling out alternative misstatement hypotheses when a nonmisstatement explanation is provided, reviewers may default to a confirmatory strategy because of the cognitive demands of the task (Klayman and Ha 1987). A confirmatory test strategy involves seeking evidence to support an hypothesis, not rule it out. Klayman and Ha (1987) argue that seeking disconfirmatory evidence, while optimal when there is a great concern for false positives, requires a great deal of mental effort and capacity. Thus, they do not believe it is at all surprising that individuals may use a confirmatory strategy when testing an hypothesis. However, Klayman and Ha (1987) argue that “there is still much to learn about how factors such as [capacity demand] and task-specific information affect hypothesis-testing strategies.”

Reviewers of substantive analytical procedure results are concerned with false positive results (i.e., Type II errors): thus seeking disconfirmatory evidence is optimal yet may not occur according to Klayman and Ha (1987). Because the review process is cognitively challenging, requiring much effort and capacity, reviewers may not be willing or able to seek disconfirmatory evidence. Thus, this dissertation must address the processing strategy used by reviewers to ensure that the choice of one or the other does not drive the results. The key control involves the evidence provided in the working papers.

To prevent processing strategy from driving any results, the working papers were designed so that reviewers should not accept the results, regardless of the processing strategy employed. In addition to not being able to rule out alternative misstatement explanations, the evidence gathered by the preparer was insufficient to support the explanation provided by the client. Thus, the reviewer should be critical of the working papers as substantive evidence because of the lack of persuasiveness of the evidence supporting the documented conclusions.

Like the other environmental and task-specific concerns described in this section, choice of processing strategy is an interesting research question itself. However, subject limitations for this dissertation prevent addressing the issue to the extent that any conclusions can be drawn. The results of the experiment, however, might shed some light on the processing strategy used by reviewers. For example, the number of alternative hypotheses listed by reviewers might indicate whether reviewers, in general, were considering alternative explanations and, if so, to what extent. Also, if the decision aid causes reviewers to list significantly more alternative hypotheses, a possible explanation is that the cue caused reviewers to switch processing strategies. However, based on the design of the working papers reviewed, this result should not result in any effect on the hypothesized results.

2.4 Summary of Chapter

This chapter describes psychological theories relevant to the working paper review process for substantive analytical procedures and introduces five hypotheses based on two external cues used to examine the theories' impacts on the review process. Taken

together, these hypotheses form a model that depicts the current review process as either helpful or detrimental, depending on the role of instruction to provide review notes, to the avoidance of committing Type II errors by reviewers. The review note process might cue counterexplanation and decrease the risk of committing Type II errors, or it might be an additional cognitive task that strains capacity and increases the risk of committing Type II errors. Regardless, the introduction of a reminder list of considerations for the review process should increase available capacity. This increase should reduce the risk of Type II errors because of the ability to consider better alternative misstatement explanations and the sufficiency of the evidence gathered.

The chapter also discusses important environmental and task-specific characteristics that need to be considered when testing the hypotheses. For this dissertation, reviewers are more likely to commit Type II errors when inheriting hypotheses for which no inconsistent information is contained in the working papers and there is no *a priori* reason to suspect that the client or preparer lacked integrity or knowledge. Also, the processing strategy used by the reviewer needs to be considered so that the results of the experiment were not dependent on the processing strategy employed.

Chapter Three

RESEARCH DESIGN

This chapter describes the research design used in this dissertation, including subjects, firm participation, the experimental task, and additional data requirements. The design is used to gather data to test the five hypotheses developed in Chapter Two. The experimental design for this study consists of a between-subjects comparison of how two external cues, instruction to provide review notes and provision of a decision aid, affect the likelihood of reviewers of analytical procedures to commit Type II errors.

Section 3.1 describes how subjects were selected for the experiment. Section 3.2 details the involvement of the two international accounting firms that participated in the experiment. Section 3.3 explains the experimental task. Section 3.4 presents the dependent construct and measures. Section 3.5 describes the additional data gathered from subjects. Section 3.6 summarizes the chapter.

3.1 Subjects

Extant studies on the working paper review process (c.f. Asare and McDaniel 1996, Phillips 1996) have used senior auditors and managers. Auditors at this level have been shown to have an objective-dominant-knowledge structure for financial statement errors (Frederick, Heiman-Hoffman, and Libby 1994; Nelson, Libby, and Bonner 1996). This structure consists of organizing financial statement errors according to management assertions about financial statements. Less experienced auditors and auditing students have been shown to organize knowledge about financial statement errors according to transactions, which is not as extensive (Frederick, Heiman-Hoffman, and Libby 1994).

Therefore, using subjects with less experience would be inappropriate because they may not have the ability to access the knowledge needed to perform the review.

Furthermore, Libby and Frederick (1990) found that seniors and managers have reasonably well-developed knowledge of both error and non-error causes of fluctuations. Thus as reviewers, these subjects should be able to generate a sufficient number of alternative misstatement explanations for the management assertions associated with underlying accounts from the analytical procedure used in the experiment.

Also, senior managers and partners are not appropriate because they rarely perform detail working paper review, even though they have the requisite knowledge structures in place to perform the task. Instead, senior managers and partners perform more general working paper reviews that focus on issues raised from working paper preparation and detailed review. Thus, subjects for this experiment will consist of both senior auditors and managers.

Although there are two levels of auditors used in this dissertation, no predictions are offered concerning differences based solely on job level because there is no training during this period concerning Type II errors.¹⁶ In addition, Type II errors and misstatements, in general, are rare events. Thus, there is no reason to believe that managers will make judgments that are less likely to result in Type II errors when

¹⁶ Neither firm that participated in this study explicitly trains managers on how to prevent Type II errors.

misstatements exist. Prior studies have found mixed results when comparing seniors and managers. Messier and Tubbs (1994) found no differences between the two groups of auditors in a belief revision experiment. However, there have been documented performance differences between senior auditors and managers in previous studies. Ramsey (1994) found that audit managers are better than audit seniors at discovering conceptual errors. Phillips (1996) found that audit managers' review strategies may impact integration of audit evidence but found no such effect for seniors. Phillips (1996) was unable to determine why this difference was found. An effect for job level is possible (and tested for in Chapter Four) because existing knowledge structures, such as taxonomic structures of transaction cycles and objectives (Frederick et al. 1994, Nelson et al. 1996, Frederick 1991) are likely to become better developed with experience. However, this knowledge is not likely to lead to a reduced propensity for committing Type II errors because of the lack of experience with and training for Type II errors.

One concern for reviewers of analytical procedures is specialized knowledge of an industry or an auditing/accounting procedure (Bonner and Lewis 1990, Libby and Luft 1993). Differences in specialized knowledge across subjects could certainly affect both the ability to generate and consider alternative misstatement hypotheses, which would then impact the likelihood of committing a Type II error.

Several controls were built into the experimental design to address specialized accounting knowledge. First, to help keep any differences from driving the results, subjects were assigned randomly across each experimental condition. Second and

perhaps more importantly, a specialized industry was selected as a setting for the experimental task (see Section 3.2 for the specific industry chosen and why). Subjects were selected based on their experience in this specialized industry. By randomly assigning reviewers with specialized knowledge of the industry across cells, there is a decreased likelihood that knowledge differences existed across experimental conditions. Although the firms involved would not allow specialized knowledge to be measured directly by asking questions about the industry (or questions relating to general knowledge), subjects did provide information on the percentage of billable audit time spent on clients and any previous work experience in the industry. These measures, which should be correlated highly with specialized knowledge, were included in covariate analyses in Chapter Four, section 4.4.2, to see if knowledge differences affected the reported results.

The study included a combined 78 subjects (22 subjects were managers, 56 subjects were seniors) from both firms. The reviewers have worked for their firm an average of 4.5 years (6.6 years, 3.9 years). On average, subjects serve 6.5 clients a year (8.4 clients, 5.8 clients), spend 46% of their billable time on clients in the specialized industry (55%, 43%), and spend 27% of their time (25%, 27%) reviewing working papers, of which 45% of the review time (44%, 45%) is for analytical procedures. All subjects spent approximately 60 minutes on the experiment, of which 18 minutes actually was spent reviewing working papers.

3.2 Firms

To help control for the effects of specialized knowledge in reviewer judgments, more **than** one international accounting firm was used in the study. Because the nature of the **experimental** task is within the area of audit quality control, firm confidentiality has been **assured** for this dissertation. However, there is no reason to believe that any other international accounting firm would perform differently on this task (unless a firm has **specific** training on preventing Type II errors), particularly if no differences are found **between** groups in this experiment. The primary reason for this expectation beyond the **lack of** specific training is that all firms must perform substantive analytical procedures in **accordance** with Statement on Auditing Standard No. 56, *Analytical Procedures* (AICPA 1988). Bloucher (1995) describes that there are only a few methods available for **effectively** performing analytical procedures. Thus, firms should not differ when **performing** or reviewing analytical procedures.

To increase the external validity of the experimental task, four key partners (**including** the National Director of Audit Quality Control, Associate Regional Director of **Accounting** and Auditing, Partner In-Charge of Audit Methodology, and a lead partner **from** an auditing innovations project) from one firm participated in the development of **the** working papers and background information to ensure that they reflect realistic **analytical** procedures performed in practice. The partner group also gave suggestions on **how to** phrase questions in the instrument in a way that subjects would best understand.

Furthermore, the partner group helped select the specialized industry, financial

serv

beli

sub:

on a

part

of th

exp

the

offic

be u

chal

instr

and

Islar

mee

subj

servic

services, to be used in the study. Financial services was selected because the partners believed it to be an industry that both uses and relies heavily upon evidence from substantive analytical procedures. Given that 45% of the subjects' review time is spent on analytical procedures, the partners' belief appears to have been supported. The partners believed that use of subjects in this industry would increase the meaningfulness of the study to their firm.

Finally, one of the members of the partner group arranged for a pilot test of the experiment at a regional office in the Midwest using some of the top experts at the firm in the financial services industry.¹⁷ The pilot test also included subjects from the local office in order to represent the subject group (senior auditors and managers) that were to be used for the actual experiment. Pilot-test subjects concurred that the task was challenging, realistic, and understandable. Few substantive changes were made to the instrument as a result of the pilot test.

The actual experiment was conducted on seven different occasions at various firm and training locations in Chicago, Cleveland, Detroit, New York, and Glen Cove (Long Island) during required financial-industry training sessions or office-wide group meetings. Also, there was no time pressure constraint employed in the study; however, subjects in different locations spent about the same amount of time to complete the case.

¹⁷ These experts were already at the location working on a firm-wide, financial services auditing innovations project.

3.3 Task Description and Administration of the Experiment

This section consists of a step-by-step walkthrough of the experiment. The experiment was conducted by the author in all but one case, which was conducted by a colleague from Michigan State University. The experiment was completed in controlled environments where subjects did not communicate with each other. Also reviewers were all provided with the same instructions, background information, and experimental materials with the exception of the experimental manipulation -- the provision of one, both, or neither external cue(s). Figure 4 illustrates the research design described in detail in the following paragraphs.

All reviewers first were presented verbally with instructions for the experiment. These instructions inform reviewers of the firm's commitment to the project and states that their responses will be independently reviewed by the accounting department at Michigan State University and that the reviewers will participate in a discussion on the results of the study. These instructions were included to motivate reviewers to give their best effort, while informing them that their specific responses will be reviewed by an outside party. Reviewers then were required to put their name on the case, which placed a level of accountability on the reviewers.

Other studies have shown that effort, measured in terms of output and cognitive Processing, increases when subjects are told that their responses will be reviewed by Partners in the firm (Anderson et al. 1994, Koonce et al. 1994). Although this approach also would have been effective for this dissertation, the firms providing subjects did not

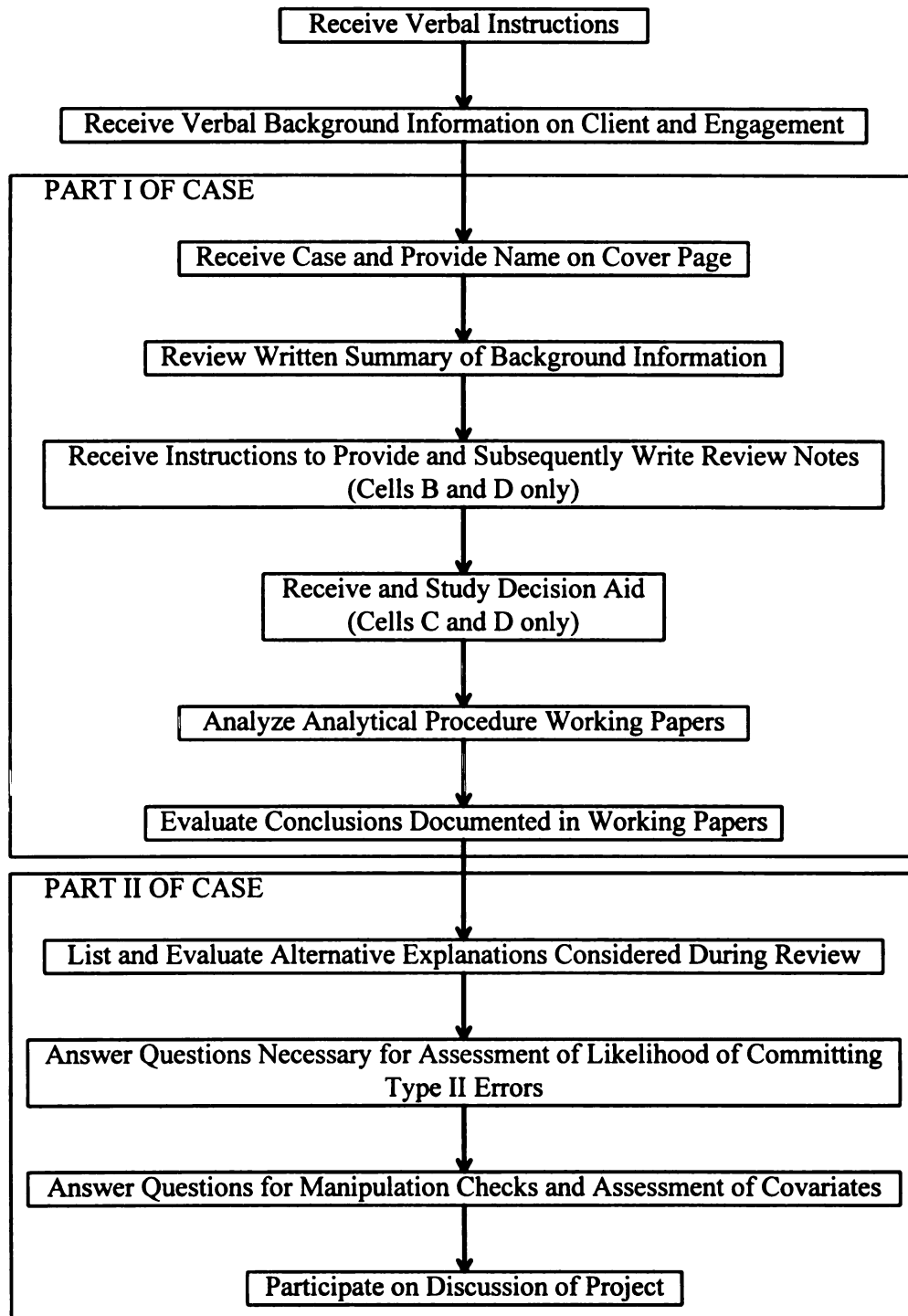


Figure 4
Research Design

want to review the completed instruments nor did they allow the use of any deceptive statements. There are other controls built into the experiment to measure effort, and they are discussed in Section 3.5 later in this chapter.

Reviewers also received instructions on how the experiment would be conducted. They were told that there were two parts of the case: the first part would involve making judgments concerning the working papers for a substantive analytical procedure in the banking industry and the second part would involve asking additional questions about the review and their personal experience. Subjects were informed that additional instructions would vary according to the version of the case received and to complete the case in the order presented. Based on results from the pilot test, subjects also were informed that there were no “right” answers or “tricks” built into the working papers. Rather, the case was designed to study the *process* of working paper review. Subjects then were informed that they would receive a verbal presentation on the background of the client and engagement, along with a written summary of the verbal presentation, which could be referred to at any time during the review.

Finally, as a control for the reviewer’s perception of what the auditing firm believes to be important for the review, subjects were informed that the primary concern when performing the review is whether material differences from expected balance fluctuations have been identified, explained, and sufficiently corroborated. Peecher (forthcoming) finds that reviewers’ perceptions of what is important to the firm when conducting a review can affect judgments. Thus, the instructions given to reviewers

prevent them from inferring that the firm is particularly concerned about performing costly procedures or gathering more evidence than is necessary to be considered sufficient. Figures 5 and 6 are the cover page and instruction page, respectively, given to each reviewer.

Reviewers are then presented verbally with background information on the client and engagement.¹⁸ This information is designed to elicit conditions under which Type II errors are likely to occur because, as previously described more fully in Chapter Two, reviewers are expected to respond to the manipulations more strongly when they have no *a priori* reason to suspect misstatements. The background information, which includes key environmental factors such as financial position and past audit issues, is the same (i.e., held constant) for subjects in all conditions. The client is described as being financially healthy, having well-maintained control systems, and having managers with high integrity. The staff auditor is described as highly competent and familiar to the

¹⁸ The verbal presentation of the background information was performed so that subjects would have a better chance to retain the information because they also were provided a written summary of the same information. Participants in the pilot study agreed that receiving the background information both verbally and in summarized, written form was helpful.

CASE STUDY**PART I****WORKING PAPER REVIEW OF ANALYTICAL PROCEDURES****Purpose Statement and Feedback Procedures***Please Read Carefully*

This case is designed to help the firm better understand key aspects of the working paper review process for substantive analytical procedures. Your effort will help to determine whether any adjustments to the current process are warranted. Every case will be independently reviewed by the accounting department at Michigan State University. There will be a feedback session at a later time. In case there are any questions during this session regarding your responses, please print your name and office location below in the space provided.

Name (Printed):

Office:

Figure 5
Cover Page of Experiment Given to Reviewers

Instructions

1. This case is designed to better understand key aspects of the working paper review process for substantive analytical procedures. Every case will be independently reviewed by the accounting department at Michigan State University. The firm supports research projects of this nature as part of its commitment to the research community. The results of this study will be considered by the firm when evaluating its current working paper review process.
2. This case contains two parts. First, you will be asked to review a set of working papers that was prepared as a substantive analytical procedure for a client in the banking industry. As a part of the review, you will be asked to make several judgments concerning the conclusions reached in the working papers. Second, you will be asked additional questions about (1) the review and (2) your professional experience.
3. Please complete the case in the order it is presented to you. Other directions and reminders appear throughout the case. These directions vary according to the version of the case you have received. Please complete the case according to the specific instructions provided. It is important for you to understand that this case is designed to study the review process. Accordingly, there are no “right” answers or “tricks” built into this case.
4. After this presentation, you will be given the first part of the case. Be sure to listen carefully to the client and engagement background information before performing the review. A summary of the background information is provided with the set of working papers. You may refer to this information at any time during the review. When completed, please raise your hand, and the first part will be collected. Then, complete the second part, which will be given to you at that time.
5. Please note that when reviewing these working papers, the primary concern is whether material differences from expected balance fluctuations have been identified and explained. It is also important to sufficiently corroborate any such explanations.

Figure 6
Instructions Given To Reviewers for Experiment

reviewer.¹⁹ Also, because the analytical procedure performed is a complex task, perceived familiarity with the preparer should better enable reviewers to identify difficulties in the preparer's conclusions (Asare and McDaniel 1996). Figure 7 is the written summary of background information provided to reviewers with Part I of the experiment.

Subjects previously had been assigned randomly to a cell manipulation. Thus, Part I of their experimental instrument included either one, both, or neither external cue(s) immediately after the written summary of the background information. Subjects in Cell A, the "control" condition, were not presented with either external cue. Subjects in Cell B, the "review note" condition, received instructions to list any review notes that would prevent them from signing off on the working papers. See Figure 8 for the exact instructions given to reviewers in the review note condition. Subjects in Cell C, the "decision aid" condition, studied the cognitive aid designed to create a retrieval plan for procedures to use when evaluating analytical procedure results (see Figure 2 in Chapter Two, Section 2.2.2, for the exact decision aid used). Subjects in Cell D, the "combined" condition, were presented with both manipulations, in a randomized order to prevent order effects.

¹⁹ Results of the pilot test indicated that the background information succeeded in conveying its intended message.

Summary of Client and Engagement Background

Financial Position

- Publicly-traded NorthBank has maintained a steady growth in interest income and earning assets
- Risk and return factors:

	<u>NorthBank</u>	<u>Peer Institutions</u>
Asset Quality (Earning Assets/Total Assets)	92.2 %	89.9 %
Liquidity (Total Loans/Total Deposits)	88.9	88.3
Interest Rate Sensitivity (Floating/Fixed Interest Rate Loans Due in 1 - 5 years)	87.7	84.5
Capital Adequacy (Primary Capital/Capital Required)	151.6	150.0
(Loan Loss Allowance/Total Loans)	1.6	1.5
(Loan Charge-Offs/Average Loans)	.4	.6
Profitability (Return on Average Assets)	1.3	1.2
Leverage (Liabilities/Equity Ratio)	7.8	8.4

Engagement History

- Accounting policies and estimates conservative
- Few audit adjustments (not isolated to any one account) in 5-year history with client
- Effective internal control structure (i.e., control risk low) and no reportable conditions in past

Other Relevant Information

- Planning materiality: \$32 million; Tolerable error/misstatement: \$16 million
- Only significant area of change through interim period: 15% growth in commercial loans receivable through third quarter
- Timing:
 - Tests of controls - Early October
 - Substantive tests of details of balances - Early October and January 5 - 17
 - Closing meeting - January 18
- Analytical procedures are primary substantive tests for revenues and expenses through the end of the year
- Staff accountant Alex Jones is reliable and ranked among top 10% in office

Your Review Task and Results of Audit To-Date

- Today is January 13, 1995
- You are reviewing gross interest income as a percentage of average earning assets
- Investigation required for significant fluctuation from prior years and/or industry average of 7.4% at October 1, 1994
- Alex Jones prepared working papers
- Key controls found to be consistently applied though interim-period testing and no reportable conditions identified
- Time budget for test, including review, is 5 hours (Alex has spent 4.5 hours)

Assume all computations are correct as stated in working papers (i.e., do not recompute!). Also, assume all references to other working papers are accurate and that Alex's interpretation of those working papers is reasonable.

Figure 7
Background Information on Client and Engagement Provided to Reviewers

Please provide below a numbered list of any review notes that you believe must be addressed in order to sign off on the working papers.

1.

Figure 8
Instructions to Provide Review Notes Given to Reviewers

Next, all subjects reviewed two experimental working papers that involved a substantive analytical procedure relating to the financial services industry. Figure 9, Panel A, was the first working paper reviewed, which is a summary of the analytical procedure results. Figure 9, Panel B was the second working paper reviewed and contained *some* detailed support for the explanation. The stated hypothesis of the client, which has been adopted by the preparer of the working papers, is that two nonmisstatement causes were responsible for the fluctuation. Specifically, the analytical procedure involved a comparison of gross interest income as a percentage of average earning assets for the current year with prior years and the current-year industry average. The client's explanations for the material decrease were that (1) average yields on investment securities decreased during the current year and (2) a shift occurred during the year from holding commercial and residential mortgage loans to holding lower-yielding commercial loans.

The working papers contained no mechanical or documentation errors, *inconsistencies* in the client explanation as documented by the preparer, or *inconsistencies* in the corroborative evidence documented by the preparer. Thus, the

PBC

	<u>Gross Interest Income</u>		<u>Average Earning Assets</u>		<u>Gross Int. Margin / Avg. Earning Assets</u>	<u>% Change from Prior Year</u>
12/31/80	P/Y \$2,192,345	/	P/Y \$20,655,987	=	P/Y 10.61%	-
12/31/81	\$2,282,451	/	\$22,351,668	=	10.21%	P/Y (3.78%)
12/31/82	\$2,267,845	/	\$24,374,122	=	9.30%	(8.88%)
12/31/83	\$1,933,658	/	\$24,510,265	=	7.89%	(15.21%)
	S-II		S-II		S-II	
12/31/84	G/L \$1,783,528	/	L-20 \$26,012,552	=	6.86%	(13.09%)
	S-II		(A) S-II		S-II	S-II

PBC - Prepared by client for auditor.

PN - Amounts traced to prior year working papers without exception.

G/L- Amounts traced to 12/31/94 general ledger without exception.

R - Recomputed without exception.

Ⓐ - Computed based on daily average (See loan working papers L-20).

⊕ - Per National Office Accounting and Auditing Group, industry average was 7.4 % as of 10/1/94 and 8.1 % as of 12/31/93. Therefore, NorthBank is 7.3 % below the industry average for 1994. Per prior year working papers, NorthBank has been within 4 % of the industry average during the past 4 years.

(b) - Per Dana Smith, Controller, decrease from last year primarily due to a decline in average yields on investment securities, residential mortgages, and consumer loans. Additionally, a shift in earning assets toward commercial loans, which are lower yielding than consumer and residential mortgage loans, negatively impacted interest income as a percentage of earning assets.

NOTE: See detailed analysis of ratio on working paper S-11.

Based on discussions with client and further auditor analysis (S-11), decrease in gross interest income as a percentage of average earning assets appears reasonable. Pass further review

Figure 9

Panel A

Analytical Procedure Working Paper Provided to Subjects for Review

(in millions)

PBC

	1994			1993			% Change
	Average Balance	Interest Income	Average Rate	Average Balance	Interest Income	Average Rate	in Average Rate
Loans							
Commercial Loans	58,873	3,558	6.20%	57,753	3,542	6.99%	(11.36%)
International Loans	897	45	5.02%	710	41	5.77%	(13.13%)
Real Estate Construction Loans	441	29	6.58%	503	35	6.96%	(5.40%)
Commercial Mortgage Loans	2,829	213	7.53%	2,368	202	8.53%	(11.74%)
Residential Mortgage Loans	2,079	169	8.13%	2,297	219	9.53%	(14.74%)
Consumer Loans	3,697	340	9.20%	3,625	362	9.99%	(7.91%)
Lease Financing	191	14	7.33%	191	17	8.90%	(17.65%)
Total Loans	19,107	1,368	7.15%	17,447	1,418	8.13%	(12.04%)
Investment Securities							
US Govt. and Agency Securities	4,540	275	6.06%	3,541	288	8.13%	(25.53%)
State and Municipal Securities	619	64	10.34%	794	82	10.33%	0.11%
Other Securities	583	30	5.42%	1,038	68	6.56%	(17.19%)
Total Investment Securities	5,712	369	6.46%	5,373	438	8.15%	(20.75%)
Interest-Bearing Deposits	814	28	3.44%	1,017	48	4.42%	(22.26%)
Federal Funds Sold and Securities for Resell	136	4	2.96%	398	15	3.76%	(21.19%)
Trading Account Securities	12	1	8.33%	78	3	3.86%	116.67%
Mortgages Held for Sale	232	15	6.47%	198	14	7.14%	(9.48%)
	328,012	31,783	8.86%	324,510	31,933	7.99%	(13.09%)

Tickmark Legend (Prepared by Auditor)

PBC - Prepared by client for auditor.

P/Y - Amounts traced to prior year working papers without exception.

G/L - Amounts traced to 12/31/94 general ledger without exception.

R - Recomputed without exception.

F - Footed without exception.

A - Computed based on daily average (See loan working papers L-25).

3b - Per Dana Smith, Controller, decrease in gross interest income as a percentage of average earning assets due to decrease in yields of investment securities, residential mortgage, and consumer loans. A comparison of these assets from 1994 to 1993 supports such an explanation:

	1994	1993	% Change
Average investment securities, residential mortgages, and consumer loans	Σ 11,488	Σ 11,295	1.7%
Interest income related to these assets	Σ 878	Σ 1,019	(13.8%)
Average Rate	7.6%	9.0%	(15.3%)

3c - Per Dana Smith, Controller, decrease in gross interest income as a percentage of average earning assets is also due to a shift in earning assets towards commercial loans, which are lower yielding than consumer and residential mortgage loans. A comparison of average earning assets from 1994 to 1993 supports such an explanation:

	1994	1993	% Change
Commercial Loans	Σ 11,802	Σ 10,121	9.9%
Other Earning Assets	Σ 14,210	Σ 14,389	(7.0%)
	Σ 26,012	Σ 24,510	
	100.0%	100.0%	

Figure 9

Panel B

Analytical Procedure Working Paper Provided to Subjects for Review

working papers were designed so that the review process was not necessary to improve the accuracy of or identify inconsistencies within the working papers.

Instead, the primary deficiency of the working papers was that the management explanations had not been corroborated fully, i.e., alternative explanations for the fluctuation had not been ruled out and the evidence gathered was not sufficient to support the given explanation reasonably. Accordingly, the only substantive working paper deficiencies that reviewers should have been able to identify were alternative misstatement hypotheses not ruled out by the preparer or the lack of sufficient evidence to support the given explanation.²⁰

Upon exposure to one, both, or neither manipulation(s) and the analytical procedure working papers, subjects completed Part I of the task by responding to the measures of the dependent construct, which are described below in Section 3.4. Subjects then handed in Part I of the experimental instrument and were given Part II, which included the final measure of the dependent construct (see Section 3.4) and additional questions relating to the experimental task and their professional experience (see Section 3.5).

²⁰ Participants in the pilot test concurred that the working papers had no inconsistencies or inaccuracies as stated. They believed failure to consider and/or rule out alternative explanations and gather sufficient evidence to support the given explanation to be the only substantive weaknesses.

3.4 Dependent Construct and Measures

The dependent construct for the working paper review process for substantive analytical procedures is the reviewer's likelihood of committing a Type II error. This construct is proxied in this dissertation by measuring the reviewer's judgment of the preparer's conclusions documented in the working papers. If the reviewer consciously is counterexplaining and considering alternative misstatement hypotheses, then documented conclusions in working papers for which alternative misstatement hypotheses have not been ruled out should not be judged as ready to be signed off on as complete (i.e., accurate and sufficiently corroborated). Furthermore, the evidence documented is insufficient to support the given explanation. This variable is used to proxy the likelihood of committing Type II errors: if a misstatement exists, the reviewer is more likely to detect the misstatement when the overall judgment of the conclusions in the working paper is lower.

This experiment is designed to test the impact of two external cues on the likelihood of committing Type II errors. Judging working-paper conclusions as complete when material alternative misstatement hypotheses have not been ruled out greatly increases the probability of the occurrence of a Type II error. Thus, the more reviewers were critical of the working papers in this experiment, the less likely they would have been to commit a Type II error.

Measuring this construct is a challenging task, particularly because a reviewer typically signals that a working paper's conclusions are complete simply by signing off

on

imp

dec

me

em

ask

crit

inv

gat

exp

anc

of t

fluc

or c

Koe

Hov

emp

que.

on the working paper itself. To maintain sufficient external validity for this study, it was important to use measures that reflect the actual decisions a reviewer makes when deciding whether to sign off on a working paper. Because the construct is difficult to measure accurately, multiple measures that theoretically measure the construct were employed.

The first four measures were questions, using a 100-point likert-type scale, that asked reviewers to evaluate the conclusions documented in the working papers using criteria that they likely use when evaluating analytical procedures. The questions involved assessing the validity, persuasiveness, sufficiency, and type of evidence gathered.

The first measure asked reviewers to evaluate the likelihood that the documented explanation accounted for substantially all (e.g., $\geq 85\%$) of the fluctuation. The scale was anchored by beliefs that “the explanations definitely did not account for substantially all of the fluctuation” (0) to “the explanations definitely accounts for substantially all of the fluctuation” (100). Previous studies involving the performance of analytical procedures or other review contexts have used similar likelihood judgments (e.g., Heiman 1990, Koonce 1992, Anderson et al. 1994, Peecher [forthcoming], and Phillips 1996). However, these studies used this question as a sole measure; whereas, this experiment employs multiple measures to reduce measurement error associated with the construct in question.

The second question is another direct indicator of the reviewer's confidence in the conclusions documented in the working papers. It asked reviewers to evaluate the extent to which additional evidence needed to be gathered before he or she was willing to sign off on the working papers. This question was anchored by beliefs that "the entire procedure needs to be reperformed" (0) to "no additional evidence needed" (100).

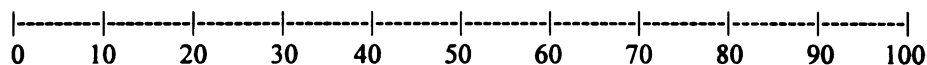
Questions three and four addressed the amount and type of evidence gathered.

The terms extent and nature were used to represent amount and type, respectively, because they are words used frequently in auditing. These questions were included because they are indicators that the explanations provided could be incorrect or that alternative misstatement hypotheses were not ruled out. The extent question was anchored by beliefs that the evidence gathered was "way too little" (0) to "just enough" (50) to "way too much" (100). The theoretical maximum for the working papers reviewed in the experiment was 50 because the explanations provided were designed to be insufficient. However, anchoring the scale only by "just enough" would create possible demand effects. Finally, the nature question was anchored by beliefs that evidence gathered was "totally irrelevant" (0) to "totally relevant" (100). Figure 10 contains the actual questions used for the first four questions in the experiment.²¹

²¹ The scales and margins are slightly modified from what subjects were provided so that the figure meets established dissertation guidelines.

Evaluation of Working Papers

1. Based on your review of the working papers and prior to the resolution of any additional issues you have with the documented explanations, what is your estimate of the probability that the documented explanation accounts for substantially all (e.g., $\geq 85\%$) of the decrease in gross interest income as a percentage of average earning assets?



**Explanation
Definitely Does
Not Account for
Substantially All
of Fluctuation**

**Unable to
Estimate Whether
Explanation Accounts
for Substantially
All of Fluctuation**

**Explanation
Definitely
Accounts for
Substantially All
of Fluctuation**

2. Assuming that all referenced working papers reflect Jones' interpretations and that other audit procedures performed *do not* disconfirm the conclusions reached by the preparer, what is your estimate of the extent to which additional corroborative evidence needs to be gathered before you are willing to sign off on the working papers?

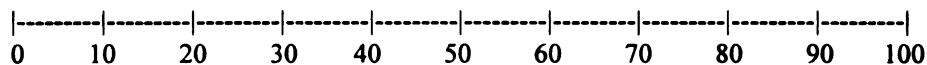


**Entire Procedure
Needs to be
Reperformed**

**Moderate Amount of
Additional Evidence
Needed**

**No Additional
Evidence
Needed**

3. Based on your answer in Number 2 above, how would you evaluate the *extent* of corroborative evidence gathered by the preparer?

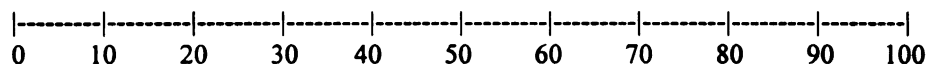


**Way Too
Little
(Ineffective)**

**Just
Enough**

**Way Too
Much
(Inefficient)**

4. Based on your answer in Number 2 above, how would you evaluate the *nature* of corroborative evidence gathered by the preparer?



**Totally
Irrelevant**

**Moderately
Relevant**

**Totally
Relevant**

Figure 10
Questions that Measure Reviewer's Judgments of Working Paper Conclusions

Data for the fifth measure was gathered immediately after collecting Part I of the instrument, which included the working papers, manipulations, and dependent measures. This measure involved asking reviewers to list all alternative explanations considered, if any, during the review and to indicate whether they believed the explanation to be ruled out by the corroborative evidence gathered. The measure that should help evaluate the reviewer's overall judgment of the documented conclusions in the working papers was the net number of alternative hypotheses believed not to have been ruled out by the evidence.

A confirmatory factor analysis analyzed the five measures to determine whether they loaded on one factor, which was designed to proxy the reviewer's overall judgment of the conclusions documented in the working papers. The details of the factor analysis are discussed in Chapter Four; however, the results indicated that the four questions loaded strongly on one factor. The fifth measure did not load on the factor, so the listing of alternative hypotheses is analyzed separately and not included as part of the dependent measure. Thus, the four questions are summed and used as a measure of the dependent construct.

Figure 11 shows a summary of the hypothesized relationships for each scenario of hypotheses described in detail in Chapter Two. These figures demonstrate how the combined measure of the reviewer's judgment of the conclusions documented in the working papers, which proxies for the likelihood of committing a Type II error, changes based on the experimental condition. If instruction to provide review notes cues

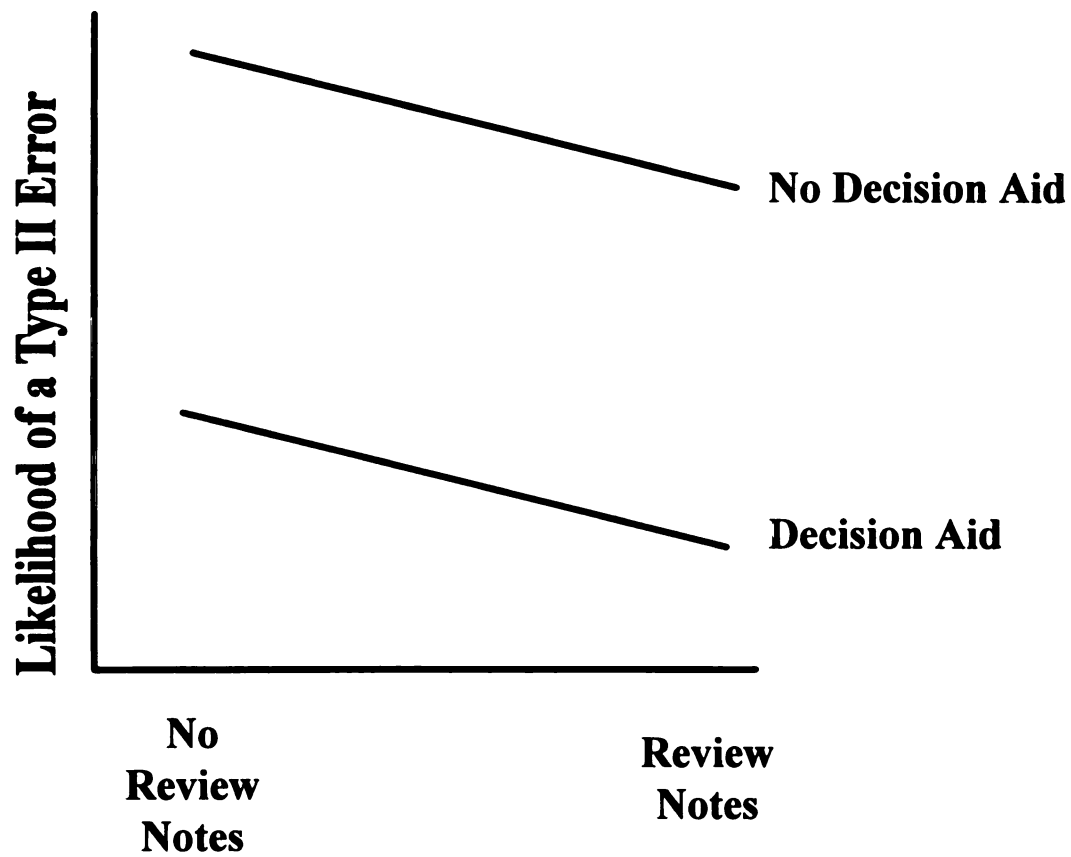


Figure 11
Hypothesized Relationships Between Dependent and Independent Variables
Panel A
Review Notes Cue Counterexplanation/Alternatives

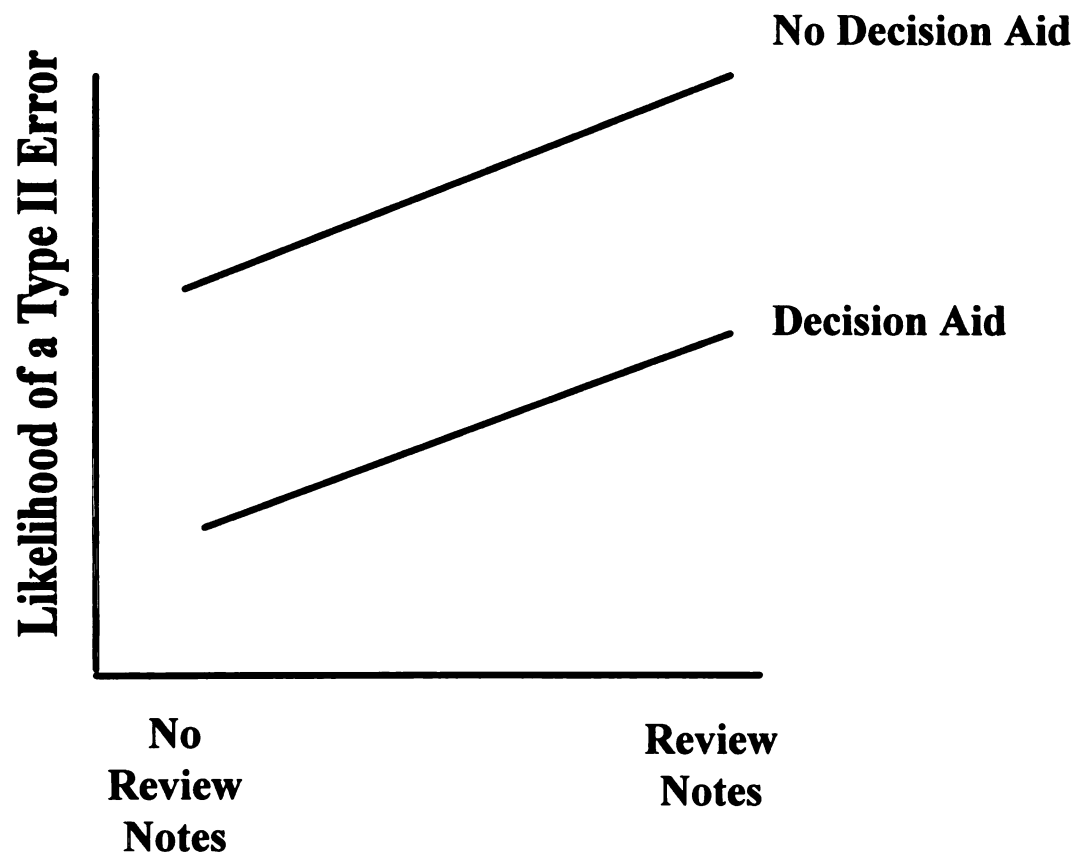


Figure 11
Hypothesized Relationships Between Dependent and Independent Variables
Panel B
Review Notes *Do Not* Cue Counterexplanation/Alternatives

counterexplanation, Panel A is the predicted relationship of the dependent measure scores and the experimental conditions. If instruction to provide review notes does not cue counterexplanation, Panel B is the predicted relationship of the dependent measure and the experimental conditions.

3.5 Additional Data Gathered

The remaining questions in Part II of the case gathered information to be used for covariate and other analyses. The issues addressed by the questions, followed by the location in the following section where the rationale for their inclusion are as follows: whether reviewers guessed on the dependent measures (Section 3.5.1); the extent that they were concerned about the effectiveness and efficiency of the analytical procedures (3.5.2); how they perceived the client and the preparer based on the background information (3.5.3); whether the manipulations were effective (3.5.4); their professional experience with the firm (3.5.5); and their effort when participating in the experiment (3.5.6). The actual questions in the experiment used to gather this additional data is included as Appendix, immediately following this dissertation.²² Also, detailed results from analyses of this data are provided in Chapter Four, Experimental Tests and Results.

²² The scales and margins are slightly modified from what subjects were provided so that the appendix meets established dissertation guidelines.

3.5.1 Reviewer Confidence

Because reviewers might have been unsure of the overall quality of the working papers, a control was included that measured the reviewer's confidence of each of the questions used to form the dependent measure. This control helped ensure that lack of confidence -- i.e., guessing -- was not driving the measures. Subjects were asked how confident they were of their likelihood assessment for each question (see Appendix, p. 152). As predicted, there generally was no correlation between the dependent measure questions and confidence ratings. This result suggests that reviewers were not guessing when answering the questions.

3.5.2 Reviewer Concern for Efficiency and Effectiveness

Peecher (forthcoming) found that it is important to control for effects of reviewers' perceptions of the auditing firm's primary concerns when performing a review. Therefore, the instructions for the instrument include language suggesting that the firm is concerned with identifying and explaining any material differences from expected balance fluctuations and sufficiently corroborating any such explanations. These instructions enable the auditing firm's primary concerns when performing a review to be consistent. Two exit questions were included to evaluate whether reviewers' judgments were impacted by their primary concerns when performing a review (see Appendix, p. 153). The results failed to show that subjects' perceptions of the audit firm's concerns were any different across experimental condition.

Based on Peecher's (forthcoming) results, an argument can be made that subjects more concerned with performing costly procedures that are not justified will be more likely to commit a Type II error. This argument is posited because reviewers who are focused on cost should be less likely to counterexplain, consider alternative explanations, or perform the level of detail review necessary to discover that insufficient evidence has been gathered to support a given explanation. Rather, the reviewer's concerns of cost are predicted to carry over to the review process as well.

Conversely, reviewers more concerned with accepting an unjustified explanation are more likely to scrutinize working papers carefully, which should include more attention to counterexplanation, alternative explanations, and detailed review of the evidence provided for sufficiency. Like reviewers who expect a misstatement, these reviewers are expected to pay less attention to the other areas of the analytical procedure (documentation, cohesiveness, etc.) when demands on capacity become excessive.

This dissertation predicted that the manipulations would be more effective when Type II errors are more likely *a priori*, and an analysis of reviewers for this experiment supports this claim. The two exit questions measured the reviewer's concern with (1) accepting client explanations without adequate justification ("effectiveness") and (2) performing costly investigations without adequate justification ("efficiency"). The difference between the two questions was used as a measure of whether the reviewer was more concerned with effectiveness $[(1) > (2)]$ or efficiency $[(1) \leq (2)]$. Differences of zero were classified as "efficiency" because these reviewers were only concerned with

effectiveness up to the point that accepting unjustified client explanations was no longer cost beneficial. Reviewers who were classified as more concerned with efficiency responded more strongly to the manipulations than those reviewers who were classified as more concerned with effectiveness. This result suggests that *a priori* likelihood of committing a Type II error was an important factor to control.

3.5.3 Reviewer Perception of Client and Preparer

Because the hypothesized results were only expected to hold when Type II errors were expected to occur, the study was designed so that the reviewer would have no *a priori* expectation of a misstatement. To determine whether perceptions of knowledge bias and reporting bias were low for the reviewers, they were asked nine questions that measured their perceptions of the background and engagement information (see Appendix, pp. 153-155).²³ The average perceptions of reviewers was above the scale's midpoint of five. However, an analysis of results suggested that only subjects who averaged a high score for the questions (above seven) had low assessments of knowledge

²³ Reporting bias and knowledge bias are discussed in detail in Chapter Two, Section 2.3.2. As a reminder reporting bias refers to the the degree to which a communicator's willingness to convey an accurate message is compromised, and knowledge bias refers to the degree to which a communicator's knowledge of message relevant information is inaccurate (Eagly and Chaiken 1993, 1984). The background information was designed to elicit low perceptions of each type of bias.

and reporting bias because only those subjects responded to the manipulations. Thus, again it appears that *a priori* likelihood to commit Type II errors was an important control for this study.

3.5.4 Manipulation Checks

As a test of the review note manipulation, subjects were asked the extent to which they provided review notes (See Appendix, p. 155). Results suggested that the review note manipulation was highly effective.

As the results in the next chapter demonstrate, the effect of using a decision aid had a robust impact on the dependent construct both when used alone and in conjunction with writing review notes. However, independent of the hypotheses test results, there is no way from the data collected to ensure that the manipulation was effective other than (1) the finding that no measured covariates were driving the results and (2) the random assignment of subjects.

3.5.5 Measures of Reviewer's Professional Experience

Demographic information was collected relating to age, sex, experience with the firm, position held for the majority of the past year, average number of engagements per year, percentage of time dedicated to reviewing working papers, percentage of time dedicated to performing analytical procedures, percentage of time dedicated to reviewing analytical procedures, percentage of time spent on clients in the financial services industry, and previous experience in the financial services industry (see Appendix, p. 158). None of these variables were found to have an impact on the results.

3.5.6 Reviewer Effort During Experiment

To check whether subjects put forth sufficient effort, a question asking subjects their effort level was included (see Appendix, p. 160). The scale ranged from much less effort than an actual review (0) to as much effort as on an actual review (5) to much more effort than on an actual review (10). Reviewers reported levels of effort that were not different across conditions (mean = 3.98).

As an additional measure of effort, a series of exit questions were included that asked subjects to what extent they considered various aspects of the working paper review process (see Appendix, pp. 156-157). Interestingly, even subjects who were not exposed to the decision aid indicated that they considered these aspects to the same extent as those subjects exposed to the aid. Thus, it appears subjects were exerting sufficient cognitive effort within and equal effort between experimental conditions. Also, the accountability manipulation seemed to have the desired effect, as subjects appeared relieved to know during the debriefing session that no firm personnel would review their individual work.²⁴

²⁴ An exit question that asked subjects to what extent they felt accountable would have been more effective. However, the firm involved in the design of the case did not allow such a question to be asked.

3.6 Summary of Chapter

This chapter describes how the experimental design was developed and implemented. Data gathered from the experiment was meaningful because subjects with the appropriate knowledge were used. These subjects--seniors from two international auditing firms--also represented reviewers who actually perform the review task used in the experiment. The experiment also was improved by the contributions of the expert partner group that participated in the design of the working papers, selection of the specialized industry, and realism of the background information.

The experiment itself was designed to place subjects in a review task of a substantial analytical procedure where they would be more likely to commit Type II errors. Then, subjects were provided one, both, or neither external cue(s) (instruction to provide review notes and provision of a decision aid). To measure the impact of the cues on the review process for substantive analytical procedures, subjects were asked to make a series of judgments that proxy for the decision process that auditors go through when deciding to sign off on a working paper. A comparison of the combined score of these measures for subjects in different cells are then used to test the hypotheses.

The next chapter describes the experimental tests and results necessary to evaluate whether this design was effective for helping to answer the research question of evaluating to what extent instruction to provide review notes and/or provision of a cognitive decision aid affects the likelihood of committing Type II errors.

Chapter Four

EXPERIMENTAL TESTS AND RESULTS

This chapter describes the experimental procedures used to test the hypotheses described in Chapter Two of the dissertation. The data used in these statistical tests was gathered from professional audit seniors and managers from two international accounting firms using the research design discussed in Chapter Three of this dissertation. This chapter also provides results of the hypotheses tests, as well as tests for manipulations, possible covariates, and other key environmental and task-related variables.

Section 4.1 describes the specification and measurement of the dependent construct. Section 4.2 discusses the classification of review notes based on whether or not they included counterexplanation or ruling out of alternative explanations. Section 4.3 describes the experimental tests of the hypotheses and discusses the results. Section 4.4 tests and discusses the effects of relevant environmental and task-specific variables. Section 4.5 describes other tests and results for prior influence and concomitant variables not discussed in earlier sections. Section 4.6 provides results for manipulation check tests. Section 4.7 provides an analysis and results of alternative explanation listings provided by subjects. Section 4.8 summarizes the chapter.

4.1 Specification and Measurement of Dependent Construct

The reviewer's likelihood for committing a Type II error is the dependent construct in question when reviewing substantive analytical procedures. For the dissertation experiment, this construct was proxied using reviewer judgment of the

quality of the conclusions documented in the experimental working papers (see Figure 9 in Chapter Three, which displays the experimental working papers). The working papers should not have been acceptable to the reviewer based on an insufficient amount of evidence gathered by the preparer. Thus, a higher overall reviewer judgment suggests an increased probability that the reviewer would have accepted the working papers and possibly committed a Type II error if a misstatement existed. Therefore, higher overall judgments for these working papers increases the *risk* of a Type II error.

To measure the proxy, four 100-point Likert-type scale questions were asked that represent the judgments reviewers make when deciding whether or not to accept a working paper as sufficiently defensible to other parties. These questions asked reviewers to assess the validity, persuasiveness, sufficiency, and type of evidence gathered (see Figure 10 in Chapter Three for specific questions asked). A fifth measure of the proxy was based on an instruction for reviewers to list all alternative explanations considered when performing the review and indicate whether the reviewer believed that the alternative explanation had been ruled out (see Appendix, p. 151, for the exact question asked). The fifth measure for this experiment was the number of alternative explanations not believed to have been ruled out less the number that the reviewer believed were ruled out. This measure was an additional indication of the reviewer's evaluation of the validity of the explanation provided by the preparer.

Prior studies involving the performance of analytical procedures or other review contexts have used a sole measure when examining an auditor's evaluation of analytical

procedure results (e.g., Heiman 1990, Anderson et al. 1994, Peecher [forthcoming], and Phillips 1996). These studies used a question similar to the first question used in this dissertation: evaluate the likelihood that the explanation being evaluated accounted for substantially all of the fluctuation.²⁵ This dissertation extends prior research by measuring additional auditor judgments concerning the persuasiveness, sufficiency, and type of evidence gathered to support the explanations provided when examining auditor judgments concerning analytical procedure results. Taken together, these measures should have provided a better indication of the reviewer's judgment of the conclusions documented in the working papers than any one measure alone. In other words, measurement error of the dependent construct should be reduced.

When a group of measured variables are hypothesized to load on one common factor (or construct), confirmatory factor analysis can be performed (Kim and Mueller 1978a). Confirmatory factor analysis tests whether hypothesized underlying factors, which are smaller in number than the observed variables, are responsible for the observed covariation among the observed variables (Kim and Mueller 1978b). For this dissertation, the five dependent measures were predicted to load on one factor, which

²⁵ Heiman (1990) asked subjects to evaluate the probability that the fluctuation was the "true" cause of the fluctuation, which is somewhat different from the other papers cited. Those papers used questions that were almost verbatim to the first question used in this dissertation.

represents the reviewer's overall judgment of the conclusions reached in the working papers. Thus, a confirmatory factor analysis was generated to test whether the measures loaded on one factor.

Results of the factor analysis showed that one of the measures, net alternative hypotheses not ruled out in the working papers, was not found to be a significant factor (factor loading $< .20$) primarily because many subjects did not list more than one net alternative hypothesis not considered to be ruled out. Thus, this measure was not included as part of the dependent construct. See section 4.7 for an analysis of subjects' listing of alternative hypotheses.

The other four dependent measures were all found to load as one factor (all factor loadings $> .70$). Figure 12 diagrams the four variables and their factor loadings on the hypothesized construct. The factor is highly reliable (Cronbach's $\alpha = .76$). The factor also explains a sufficient amount of variance in the data gathered for the dependent measure (eigenvalue = 2.31 and percentage of variance explained [eigenvalue / number of variables] = .58). Means and standard deviations for dependent measures in each condition are presented in Table 1, along with factor scores for each measure. Table 1 also includes means and standard deviations for the dependent factor, which is measured as the sum of each of the dependent measures included in the factor. This factor should represent the reviewer's judgment of the quality of the conclusions reached by the preparer in the working papers. Figure 13 plots the means for each dependent measure shown in Table 1.

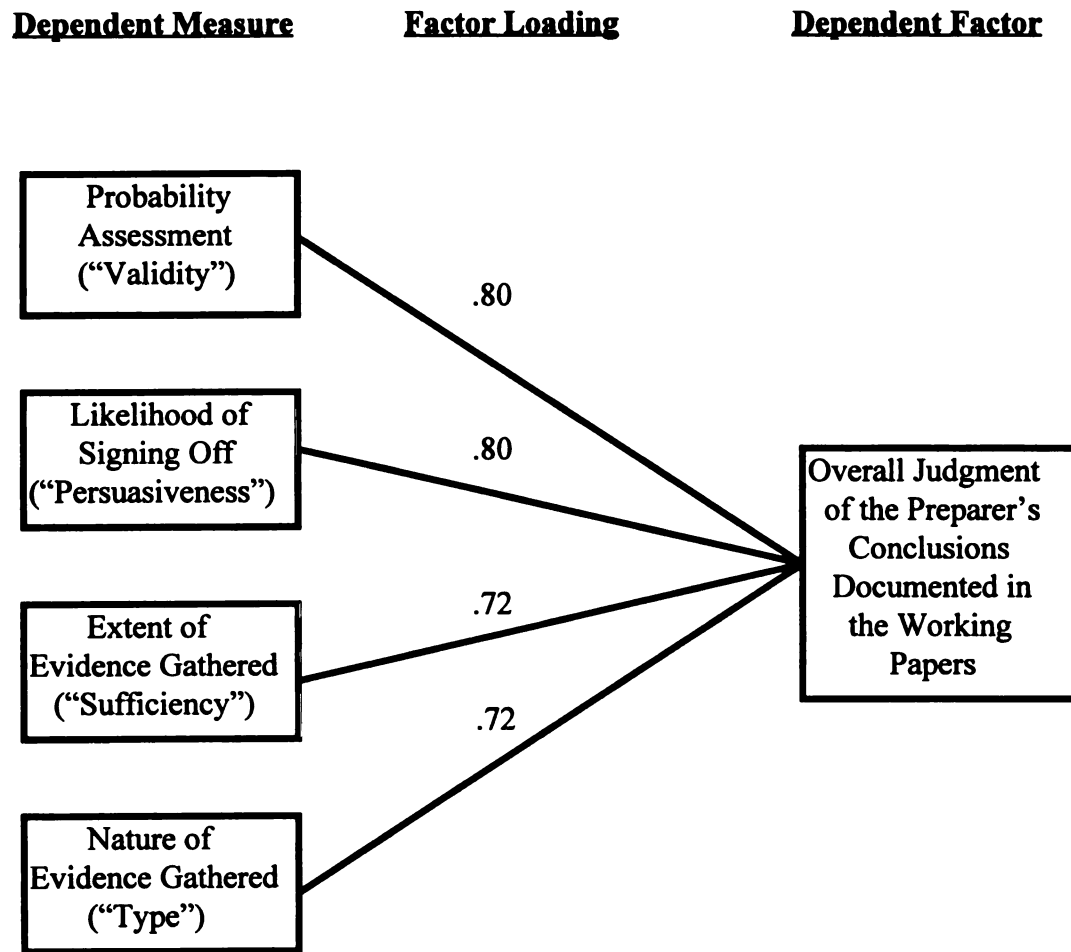


Figure 12
Diagram of Confirmatory Factor Analysis

Table 1
Specification and Descriptive Statistics for Dependent Construct

Dependent Measure [Factor Loading]	Independent Manipulation Condition			
	Cell A	Cell B	Cell C	Cell D
	No Review Notes No Decision Aid (n = 22)	Review Notes No Decision Aid (n = 19)	No Review Notes Decision Aid (n = 18)	Review Notes Decision Aid (n = 19)
Probability Assessment [.80]	60.18 (22.50) Mean (Std. Dev.)	62.37 (22.81)	50.61 (30.56)	65.05 (20.29)
Likelihood of Signing Off [.80]	58.50 (22.30)	66.58 (17.24)	52.22 (23.21)	62.37 (20.97)
Extent of Evidence Gathered [.72]	41.32 (12.02)	44.63 (12.03)	37.00 (12.83)	38.68 (13.10)
Nature of Evidence Gathered [.72]	69.00 (19.08)	65.00 (24.09)	59.11 (23.58)	65.79 (19.53)
Dependent Factor*	229.00 (53.62)	238.58 (74.55)	198.94 (74.55)	231.90 (57.20)

Factor Analysis Data:

Eigenvalue: 2.312

Variance Explained: .58

Cronbach's Alpha: .76

* - Dependent factor is sum of four measures for each subject. The factor can be interpreted as each subject's overall judgment of the preparer's conclusions documented in the working papers. The measures summed were all included because their factor loadings were above .4.

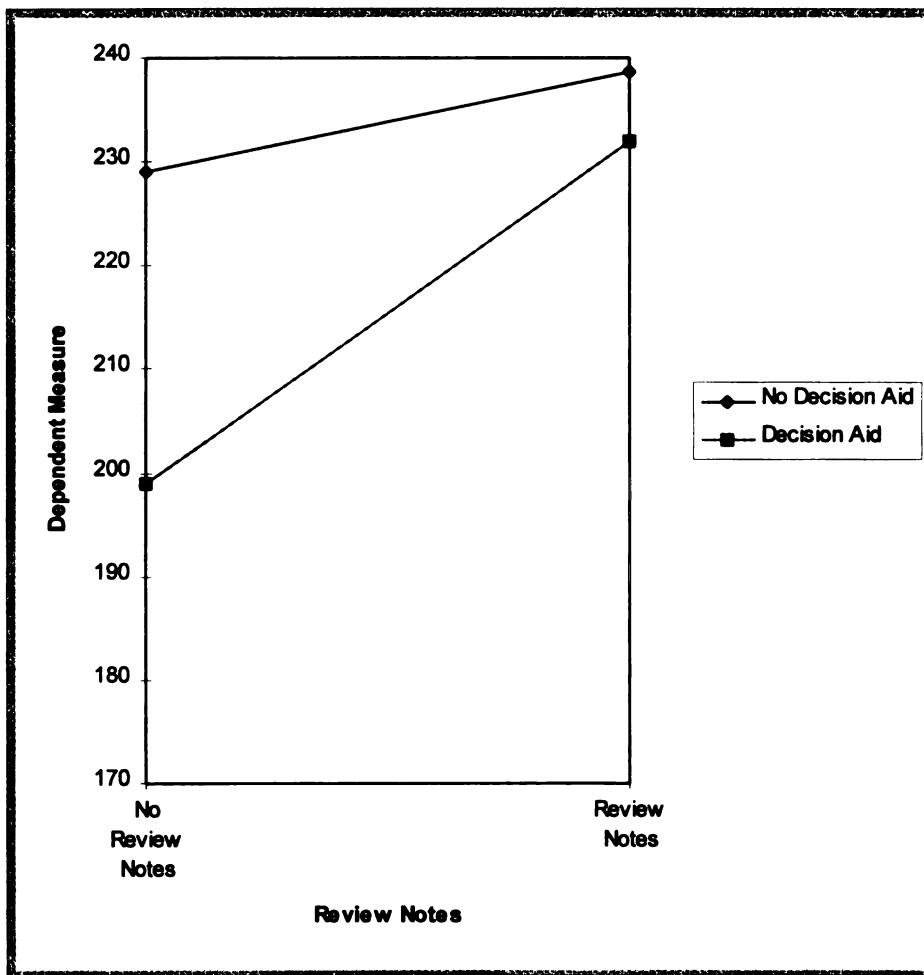
4.2 Classification of Review Notes

An argument can also be put forth that inclusion of counterexplanation or alternative hypotheses in the review notes indicated that instruction to provide review notes cued counterexplanation or the generation of alternative explanations. Failure to include such items also can be argued to be an indication that the instruction to provide review notes did not cue counterexplanation or generation of alternative explanations. As a means of testing Hypotheses 1 and 5, review notes needed to be classified according to whether they contained counterexplanation or alternative explanations.

Two independent judges with public accounting experience classified subjects' review notes according to whether they contained explicit counterexplanation of the preparer's hypothesis or alternative hypotheses not ruled out.²⁶ The classifiers initially agreed on 34 out of 38 classifications, and they subsequently met and resolved the 4 differences.²⁷ They classified 4 of 19 (21%) and 3 of 19 (16%) review note lists in the review note and combined conditions, respectively, (or 7 of 38 [18%] in total) as

²⁶ Both independent judges were fourth-year doctoral students. One student had four years of professional auditing experience and the other had one year of professional auditing experience.

²⁷ The 38 classifications represent the 19 reviewers in the review note condition and the 19 reviewers in the combined condition. These reviewers were the only subjects instructed to provide review notes.



This figure is a plot of actual cell means for the dependent measure, which can be interpreted as each subject's overall judgment of the preparer's conclusions documented in the working papers. The corresponding cells are as follows:

Control Condition:	No Review Notes / No Decision Aid
Review Note Condition:	Review Notes / No Decision Aid
Decision Aid Condition:	No Review Notes / Decision Aid
Combined Condition:	Review Notes / Decision Aid

Figure 13
Plot of Means for Dependent Measure

containing either counterexplanation or alternative hypotheses. The review notes focused on documentation issues and requests to perform more procedures to *support* the explanations in the working papers, which suggests that reviewers were following a confirmatory processing strategy (see Chapter Two, section 2.3.3 for a discussion of processing strategy). These results suggest that instruction to provide review does not cue counterexplanation possibly because of a strain placed on available capacity. This argument is made given that subjects in the combined condition were provided with a decision aid that included instructions to consider alternative explanations not ruled out in the working papers, yet they were no better than subjects in the review note condition at including counterexplanation or alternative explanations.

A comparison of the dependent construct scores across review note classifications further suggests that instruction to provide review notes does not cue counterexplanation. Means for dependent measures in the review note only condition were not different (240.75 versus 238) for subjects classified as providing counterexplanation or alternative hypotheses versus not providing either. Furthermore, the presence of *some* counterexplanation or alternative hypotheses may not be enough. The mean of 240.75 for those who included counterexplanation or alternative hypotheses still was higher than the mean of 229.00 for subjects in the control condition, who received no external manipulation (no significance test was performed because only 4 subjects provided counterexplanation).

Also, no other factors were found that drove reviewers to be more or less likely to include counterexplanation or generate alternative explanations. A check of factors such as job level, firm experience, percentage of billable time spent reviewing working papers or on financial services clients, time spent reviewing working papers in this experiment, and self-reported effort on the experimental instrument failed to indicate that any characteristic had a significant effect on those who provided counterexplanation or alternative hypotheses.

Thus, for the analyses in this study, all subjects' scores in the review note condition are combined and only Hypothesis 1b is tested because it appears that instructions to provide review notes does not cue counterexplanation or generation of alternative explanations (there is no effect on the results when removing the subjects who included some counterexplanation or alternative hypotheses). Also, Hypothesis 5b is the appropriate test based on the impact of review notes when no decision aid is present.

4.3 Tests of Hypotheses

Ordinarily, a 2 x 2 ANOVA (review note x decision aid manipulation) would have been used to evaluate whether review notes and/or the decision aid had an impact on reviewer judgments concerning the conclusions reached by preparers in the working papers. However, there is a nonsymmetrical, ordinal relationship predicted when the factors for Hypotheses 1b, 2, 3, 4, and 5b (the applicable hypotheses given that review notes were not found to proxy for counterexplanation or alternative explanations) are combined into one model (see Figures 11, panel B, and 13 for the hypothesized and

actual relationship between the factors).²⁸ In addition, the slopes for each main effect are not known, *a priori*. ANOVA is a less powerful statistical tool for these types of relationships (Buckless and Ravenscroft 1990, Keppel 1982).

Instead of an ANOVA, contrast coding is selected as a more powerful method for testing the relationship suggested by the five pair-wise comparisons suggested by Hypotheses 1b, 2, 3, 4, and 5b. Contrast coding involves the selection of orthogonal weights, which must sum to zero, that are applied to each cell to enable a test of an *a priori* hypothesized model (Buckless and Ravenscroft 1990). Assigning the terms Cell A, B, C, and D to the control, review note, decision aid, and combined conditions, respectively, the five hypotheses can be written as follows:

H _{1b} :	Cell B	>	Cell A
H ₂ :	Cell C	<	Cell A
H ₃ :	Cell C	<	Cell B
H ₄ :	Cell D	<	Cell B
H _{5b} :	Cell D	>	Cell C

The hypothesized model shown in Figures 11, Panel B, and 13 is a figurative representation of the relationships shown above. The appropriate contrast weights to assign to Cells A, B, C, and D are 0, 1, -1, and 0, respectively.²⁹ Analyzing this model

²⁸ See Buckless and Ravenscroft (1990, pp. 934 -38) for a comparison of disordinal and ordinal interactions.

²⁹ The disordinal weights assigned by ANOVA are 1, -1, -1, and 1 for Cells A, B, C, and D, respectively, which is not hypothesized for the cells in this dissertation.

simultaneously tests all hypotheses (Cells A and D are equal because no hypothesis is offered for this comparison). Table 2, panel A, presents the means and contrast weights for each cell. A contrast coding analysis demonstrated that the contrast was significant ($F = 3.96, p = .025$).³⁰ Table 2, panel B, displays the results of the contrast coding analysis.

The test results support the hypothesized relationships; however, a detailed look at the relationships between the cells provides possible insights into how the external cues affected the reviewers. Clearly, the introduction of a decision aid had a significant effect on the reviewers. Those reviewers in the decision aid only cell had the lowest mean (198.94). Thus, these reviewers were the most critical of the working papers. Table 1 shows that reviewers in this cell were more critical than reviewers in every other cell for each measure of the dependent factor. As discussed in more detail in Chapter Two, this result suggests that the decision aid relieves demands on reviewer capacity that are placed by the review process by providing reviewers with a retrieval plan to use when evaluating substantive analytical procedures (particularly when comparing the decision aid and control condition).

³⁰ All p-values are one-tailed because the hypothesized relationships are one-directional.

Table 2
Impact of Review Notes and/or Decision Aids on Review Judgments

Panel A: Cell Means and Contrast Weights

Means
 (Standard Deviations)
 [Contrast Weights]

	No Instructions to List Review Notes	Instructions to List Review Notes
No Decision Aid	229.00 (53.62) [0]	238.58 (56.64) [+1]
Decision Aid	198.94 (74.55) [-1]	231.90 (57.20) [0]

Panel B: Contrast Coding Analysis

<u>Factor</u>	<u>Sum of Squares</u>	<u>df</u>	<u>Mean Square</u>	<u>F</u>	<u>p-value</u>
Model Contrast	14,529.42	1	14,529.42	3.96	.025
Residual	3,431.06	2	1,715.53	.47	> .10
Error	<u>271,509.37</u>	<u>74</u>	3,669.05		
	289,469.85	77			

Probably the most informative result was the difference between the decision aid condition and the combined condition. This comparison, which was described in detail in the Chapter Two discussion of Hypothesis 5b, provides insight on the additional demands on capacity placed on reviewers by providing them the instruction to provide review notes. Given that the decision aid was designed to create a retrieval plan to aid capacity demands on reviewers, instruction to provide review notes should not affect reviewers who have been provided a decision aid if the process of writing review notes is not an additional cognitive task requiring cognitive capacity. Also, if instruction to provide review notes cues counterexplanation, then the combined condition should have been more critical of the working papers than the decision aid condition.

However, the only reason that reviewers in the combined condition would be less critical than reviewers in the decision aid condition is if the review note process is an additional cognitive task requiring effort and thus cognitive capacity. One reason that this explanation is likely was that the working papers were designed to be deficient regardless of whether a disconfirmatory or confirmatory processing strategy was chosen. Thus, the decision aid should have resulted in reviewers in the combined condition to be at least as critical as the reviewers in the decision aid condition. Instead, the entire effect of the decision aid (control condition versus decision aid condition) was offset by the inclusion of instruction to provide review notes (decision aid versus combined condition).

The comparison of the decision aid condition with the combined condition also helps rule out demand effects as a reason for the effect of the low mean in the decision

aid cell. If the decision aid created an increased level of suspicion in reviewers by its mere presence, then the same effect should have occurred for subjects in the combined condition. In fact, because instruction to provide review notes is part of current practice, reviewers in the combined condition should have been even more likely to view the decision aid as some type of warning. However, the absence of increased skepticism of reviewers in the combined condition versus the control condition suggests that the decision aid did not produce a demand effect.

Although the contrast was significant, the decision aid only condition mean and the directional relationships of the other condition means are the main reasons; whereas, the comparisons between the review note and control conditions (H1b) and the review note and combined conditions (H4) are not statistically significant when tested outside the model using t-tests.³¹ There are several possible reasons for the lack of a significant difference. First, there could be other variables that affected the dependent measure, which can be pulled out with subsequent analysis. The following sections describe the results of such analysis, which demonstrate that other factors can be identified to improve the results. Limitations on the number of subjects, however, prevent the ability to parcel out at the same time all factors found to influence the results to some degree. Also, the

³¹ T-test statistics for each comparison were .58 ($p > .10$) for the review note versus control condition and .35 ($p > .10$) for the review note versus the combined condition.

possibility exists that the lack of a structured review process creates a significant strain on capacity in the control condition that, while worse in the review note condition, is affected by a “ceiling effect.” In other words, reviewers in the review note condition might have encountered difficulty considering alternative explanations or counterexplaining given the instruction to provide review notes, but reviewers still would have encountered the same or slightly less strain had they been in the control condition. Simply put, the review process could be a cognitively challenging process regardless of whether review notes are required when no cognitive guidance is provided. However, given cognitive guidance in the form of a decision aid, the cognitive strain is relieved *unless* there is the additional cognitive requirement to provide review notes.

To demonstrate the lack of a main effect for the instruction to provide review notes but a presence of an interaction between this instruction and the provision of a decision aid, a contrast model using the weights 1, 1, -3, and 1 for the control, review note, decision aid, and combined conditions, respectively can be performed. This “magic cell” contrast model, while not initially hypothesized, appears to be the best fit for the data. The model contrast explains more variance than the hypothesized model ($F = 4.41$; $p < .02$), while the unexplained variance from the contrast model is highly insignificant ($F = .24$). Thus, while this unhypothesized model might have occurred by chance (i.e., sampling error), the likelihood exists that provision of a decision aid relieves strain on capacity unless the instruction to provide review notes is also given. Table 3 shows the contrast coding analysis for this model.

Table 3
Contrast Coding Analysis for “Magic Cell” Model

Panel A: Cell Means and Contrast Weights

Means
 (Standard Deviations)
 [Contrast Weights]

	No Instructions to List Review Notes	Instructions to List Review Notes
No Decision Aid	229.00 (53.62) [+1]	238.58 (56.64) [+1]
Decision Aid	198.94 (74.55) [-3]	231.90 (57.20) [+1]

Panel B: Contrast Coding Analysis

<u>Factor</u>	<u>Sum of Squares</u>	<u>df</u>	<u>Mean Square</u>	<u>F</u>	<u>p-value</u>
Model Contrast	16,180.51	1	16,180.51	4.41	.019
Residual	1,779.97	2	889.99	.24	> .10
Error	<u>271,509.37</u>	<u>74</u>	3,669.05		
	289,469.85	77			

This dissertation only identified the question and provided some initial answers. Future research on the review note process will be necessary to further clarify the exact cognitive role of the review note process. Clearly, instruction to provide review notes does not cue counterexplanation or the generation of alternative explanations. However, at this point, the cognitive demands of the review process, absent instruction to provide review notes, is not understood well enough to answer fully the question regarding the effect of what appears to be additional cognitive effort needed to consider and write review notes.

Because of the complexity of the review task and the results previously discussed, tests are needed to see whether the manipulations were more or less effective under certain conditions. The remainder of this chapter focuses on reviewers' perceptions of the client and engagement, effectiveness and efficiency prioritization, effort and accountability, specific experience, responses to manipulation checks, and listings of alternative explanations. Taken together, the results of the hypotheses tests are shown to have been strongest under conditions in which reviewers were likely to commit a Type II error.

4.4 Tests for Environmental and Task-Specific Variables

The following section describes results of tests of the contrast model after controlling for other key environmental and task-specific variables. The factors that were measured were considered to be possible influences of reviewer behavior that needed to be addressed. Section 4.4.1 examines *a priori* likelihood to commit Type II errors, which

was discussed in detail in Chapter Two. This factor was measured and tested because the manipulations were hypothesized to affect reviewers as described in this dissertation when they were likely to commit Type II errors. Section 4.4.2 examines reviewer motivation, which was addressed in terms of effort and accountability. Given the complexity of the task, sufficient effort in terms of information processing is necessary before the capacity overload argument can be defended.

4.4.1 *A Priori* Likelihood of Committing Type II Errors

Because the hypothesized results were only expected to hold when Type II errors were likely to exist, this study held both reporting and knowledge bias constant at low levels.³² These conditions should exist in many working paper review situations.³³ To determine whether actual reviewer perceptions of reporting and knowledge bias were

³² Reporting bias and knowledge bias are discussed in detail in Chapter Two, Section 2.3.2. As a reminder, reporting bias refers to the degree to which a communicator's willingness to convey an accurate message is compromised, and knowledge bias refers to the degree to which a communicator's knowledge of message relevant information is inaccurate (Eagly and Chaiken 1995, 1984).

³³ An example in which reporting bias would be moderate or high is when a client is known to have lied in the past. An example in which knowledge bias would be moderate or high is when a reviewer is working with a new staff auditor.

low, an analysis was performed of nine questions that measured their perceptions of the background and engagement information (see Appendix, pp. 153-155).

The mean and median of subjects' average perceptions of the questions relating to the client and engagement was 7 (on a 1 to 10 scale), which was significantly greater than the middle score of 5 ($t = 20.23$, $p < .0001$).³⁴ Also, no mean for any individual question was less than 5. Thus, subjects appeared to exhibit fairly low expectations of reporting bias. Subjects evaluated the preparer as being highly competent (a mean of 7.85 on a 1 to 10 scale) and they indicated that they normally work with above average staff workers (a mean of 6.86 on a 1 to 10 scale). Thus, subjects appeared to exhibit fairly low expectations of knowledge bias.

However, to further examine whether subjects were likely to commit Type II errors, they were split into two groups at the median of seven and the hypothesized model was tested for each group. Results suggest that those whose average perceptions were less than seven may not have had low enough perceptions of knowledge or reporting bias. Testing the hypothesized model using the 43 subjects whose average perception scores were at least 7 resulted in support for the model ($F = 4.11$; $p < .025$). However, the

³⁴ The first question, which concerned audit risk, was reversed scored so that a 1 represented high reporting bias and a 10 represented low reporting bias. All other questions were set up so that reversing scores would not be necessary -- a 1 represented high reporting or knowledge bias and a 10 represented low reporting or knowledge bias.

means for the 4 cells were equal ($F = 0.00$; $p < .96$) for the 34 subjects who had mean average perceptions less than 7.³⁵ Clearly, this result suggests that the manipulations only have their predicted effects when reporting or knowledge bias was sufficiently low.

As discussed in detail in Chapter Three, Section 3.5.2, Peecher (forthcoming) found that controlling for the effects of reviewers' perceptions of the auditing firm's primary concerns when performing a review is important. Thus, two questions in Part II of the experiment measured the reviewer's concern with (1) accepting client explanations without adequate justification ("effectiveness") and (2) performing costly investigations without adequate justification ("efficiency"). The difference between the two questions was used as a measure of whether the reviewer was more concerned with effectiveness [$(1) > (2)$] or efficiency [$(1) \leq (2)$]. Differences of zero were classified as "efficiency" because these reviewers were only concerned with effectiveness up to the point that accepting unjustified client explanations was no longer cost beneficial.

Application of the hypothesized contrasts to the 47 subjects who were more concerned with efficiency was highly significant ($F = 8.77$; $p < .003$). There was no difference between cell means for the 30 subjects more concerned with accepting client responses without adequate justification ($F = .01$; $p < .92$).

³⁵ One reviewer did not complete Part II of the experiment. Thus, a total of 77 subjects are included for the additional tests. Removing that one subject from the test of the dependent measure had no effect on the results reported in Table 2.

Thus, reviewers who were more concerned with audit efficiency than effectiveness responded to the manipulations. Given that the study was designed to elicit low assessments of reporting and knowledge bias of the client and the preparer, respectively, it follows that many subjects were more concerned with efficiency because the client does not appear to pose a great deal of risk. In other words, the reviewer's concerns of cost are predicted to carry over to the review process. However, those auditors more concerned with effectiveness in relation to cost are not as susceptible to committing a Type II error because they are more likely to perform the procedures contained in the decision aid without the need of an external cue. Like reviewers who expect a misstatement, these reviewers are expected to pay less attention to the other areas of the analytical procedure (documentation, cohesiveness, etc.) when demands on capacity become excessive. Table 3 provides a summary of results for the tests of variables that were hypothesized to be indicative of an *a priori* likelihood to commit a Type II error.

4.4.2 Reviewer Motivation

Recent research by Anderson et al. (1994) and Koonce et al. (1994) demonstrate that preparers put forth more effort when they anticipate the review process. The accountability manipulation in these previous studies was informing subjects that their work would be reviewed by a partner and requiring them to put their name on their case materials. The researchers' results suggest that accountability makes a difference in effort, measured in terms of output and cognitive processing strategies, put forth by

Table 4**Tests for *A Priori* Likelihood of Committing Type II Errors**

	<u>Cell Means</u>				<u>Contrast Model</u>	
	<u>Control Condition</u>	<u>Review Note Condition</u>	<u>Decision Aid Condition</u>	<u>Combined Condition</u>	<u>F</u>	<u>p-value</u>
Average Client/Preparer Perception Score $\geq 7^+$ (n = 43)*	233.36	248.18	194.20	234.55	4.11	p < .025
Score < 7 (n = 34)	225.10	225.37	223.57	228.25	0.00	p < .960
More Concerned with: Efficiency $^{\Delta}$ (n=47)	242.85	250.00	177.91	239.36	8.77	p < .003
Effectiveness (n=30)	208.00	228.30	232.00	211.00	.01	p < .92

+ Reviewers with average perception scores of at least seven were hypothesized to be more likely to commit a Type II error.

* One reviewer did not complete Part II of the experiment. Thus, a total of 77 subjects were included. Removing that reviewer from the test of the dependent measure had no effect on the results given in Table 2.

Δ Reviewers more concerned with efficiency were hypothesized to more likely to commit a Type II error.

subjects in behavioral auditing studies. Accordingly, as part of this experiment, reviewers were required to put their name on their instruments and were informed that they would participate in a group discussion after a brief review of the responses.³⁶

To check whether subjects put forth sufficient effort, a question asking subjects their effort level was included (see Appendix, p. 160). The scale ranged from much less effort than an actual review (0) to as much effort as on an actual review (5) to much more effort than on an actual review (10). A total of 67 reviewers indicated an effort level of at least 3 (only 2 scores under 2), and reported levels of effort were not significantly different across conditions. Also, the accountability manipulation seemed to have the desired effect, as subjects appeared relieved to know during the debriefing session that no firm personnel would review their individual work.³⁷

³⁶ This discussion occurred as part of the reviewers' industry training session. Also, the instructions stated that their work would be reviewed by the accounting department at Michigan State University, which was an additional, albeit weaker, control for accountability. Although having firm partners review subjects' responses also would have been effective, the firms providing subjects did not want to review the completed instruments nor did they allow the use of any deceptive statements.

³⁷ Here again, an exit question that asked subjects to what extent they felt accountable would have been more effective. However, the firm involved in the design of the case did not allow such a question to be asked.

As an additional measure of effort, a series of five questions were asked in Part II that measured to what extent reviewers considered various aspects of the working paper review process (see Appendix, pp. 156-157). These questions contained information that was included in the decision aid. Interestingly, even subjects who were not exposed to the aid indicated that they considered these aspects to the same extent as those reviewers exposed to the aid. Table 4 provides condition means for each question and the results of a 2 x 2 (review note by decision aid manipulation) ANOVA generated for each question in which the question response mean was used as the dependent measure. There was no significant effect of experimental condition for any of the five processing questions. Also, no mean for any question was below five, except for the question asking whether any abnormal considerations were given to the review. These results suggest that subjects were exerting a significant amount of cognitive effort within and equal effort across experimental conditions.

The results of these tests for effort also identified an interesting research question for a future study. Clearly, the decision aid had an effect on reviewers. Regardless of how the reviewers are segregated, those reviewers in the decision aid only cell who were predicted to respond to the manipulations consistently were the most skeptical (i.e., their scores for the dependent measure were significantly lower than the other conditions). However, as shown in Table 4, reviewers reported that they considered the items in the decision aid to the same extent. While this result helps rule out that the decision aid

Table 5**Comparison of Conscious Processing Reported by Reviewers**

Self-Reported Processing Question ^Π	Cell Means				ANOVA Results ^Ω (Review Note Effect) [Decision Aid Effect]	
	Control Condition	Review Note Condition	Decision Aid Condition	Combined Condition	E	p-value
Step-by-Step Process	5.42	5.37	5.18	4.58	(.32) [.81]	(p < .57) [p < .37]
Abnormal Considerations	3.16	2.86	2.94	3.68	(.22) [.42]	(p < .64) [p < .52]
Alternative Explanations	5.93	5.17	5.81	5.97	(.35) [.45]	(p < .56) [p < .50]
Quantitative Support	7.21	6.26	6.67	7.13	(.42) [.19]	(p < .52) [p < .67]
Client/Preparer Riskiness	5.56	4.76	5.40	5.24	(.78) [.08]	(p < .38) [p < .78]

Π Each question asked reviewer to respond by evaluating the extent to which each component of the review process was considered when performing the working paper review in the experiment. The response was on a 1 (None) to 10 (Much) likert-type scale.

Ω The ANOVA generated for each processing question was a 2x2 (review note by decision aid manipulation) design. The F-statistic and p-values are given for each main effect. There were also no significant interactions for any question. Thus, the responses were not different between conditions.

provided any new information to reviewers, it creates the question of why all reviewers believed they considered each item to the same extent.³⁸ The aid was designed to create a retrieval plan to ease demands on capacity. Thus, the question can be asked: does an individual know that capacity has been reached, and is it possible to know that a cue has helped to relieve the demands on capacity? To answer this question, a within-subject experiment may be necessary.

4.5 Prior Influence and Concomitant Variables

4.5.1 Reviewer Confidence

It is possible that reviewers might not be sure of the overall quality of the working papers. Therefore, reviewers may provide estimates that are not based on impacts from the manipulations. To ensure that lack of confidence is not driving the measures, subjects also were asked on a 0 (not at all confident) to 100 (extremely confident)-point scale how confident they were of their likelihood assessments for each question used for the dependent construct (see Appendix p. 152). Mean confidence scores for each of the dependent measures were 69, 75, 72, and 75 for the validity, persuasiveness, sufficiency, and type of evidence question, respectively, suggesting that subjects were fairly confident for each dependent measure response. Also, there were no significant differences in

³⁸ Had new information been included in the decision aid, it is highly unlikely that reviewers who were not provided the aid would have believed they considered the information to the same extent as those reviewers who were provided the aid.

confidence levels across conditions. As predicted, there generally was no correlation between the dependent measures collapsed across conditions and confidence ratings, except for the evaluation of the sufficiency of evidence gathered question (Pearson $r = .26$, $p < .03$). Thus, subjects were not guessing more in any one condition, if they were guessing at all. In fact, given that alternative error hypotheses were not ruled out and that the provided explanation was not corroborated at a detailed level, subjects could have been overconfident in their evaluations. Further, reviewers perceived themselves approximately equal to other reviewers at the same level in the firm, which provides further evidence that a lack of confidence was not driving the results (difference in ratings of $-.13$; $t = -.54$, $p < .82$).³⁹

4.5.2 Measures of Reviewers' Professional Experience

Demographic information was also collected relating to age, sex, experience with the firm, position held for the majority of the past year, average number of engagements per year, percentage of time dedicated to reviewing working papers, percentage of time dedicated to performing analytical procedures, percentage of review time spent reviewing

³⁹ This difference was based on the difference between scores on two questions in Part II of the instrument. The first question asked reviewers to rate their ability as reviewers. The second question asked reviewers to rate the ability of others at the firm at the same level as reviewers. The difference is the response of the first question less the response of the second question (see Appendix, p. 159 for the specific questions).

Table 6**Professional Experience Descriptive Data**

<u>Variable</u>	<u>Cell Means</u>				<u>Grand Mean</u> ^Σ
	<u>Control Condition</u>	<u>Review Note Condition</u>	<u>Decision Aid Condition</u>	<u>Combined Condition</u>	
Age in years	30	26	28	29	28
Years of Firm Experience	4.4	4.1	4.5	5.3	4.6
Years Experience at the Same Level	1.8	1.6	1.7	2.2	1.8
Number of Clients Per Year	7.7	5.9	7.6	4.9	6.5
% of Time Spent on Working Paper Review	27.0	26.9	24.7	27.4	26.6
% of Time Spent on Analytical Procedure Preparation	19.7	22.2	17.6	16.2	18.9
% of Time Spent on Analytical Procedure Review	13.8	12.9	10.4	10.0	11.9
% of Time Spent on Financial Services Clients	49.3	37.5	47.1	49.6	46.1

Σ No variables were significantly different across experimental conditions.

analytical procedures, and percentage of time spent on clients within the applicable industry. Table 5 provides condition means for each experience variable.

These professional experience variables were tested to assess possible covariates, and there were no factors that had an influence on the results beyond what has already been stated elsewhere in the paper. Of particular interest is that none of the various experience indicators had a significant impact on the results. Thus, all subjects were combined for all analyses. However, due to subject availability, 56 of the 78 (72%) reviewers were seniors. Therefore, the possibility exists that more subjects are needed at the manager level to rule out effects due to knowledge gained through experience; however, there were roughly the same number of more senior reviewers in each cell.

4.6 Manipulation Checks

As a test of the review note manipulation, subjects were asked, on a scale from 1 (no review notes provided) to 5 (some review notes provided) to 10 (many review notes provided), to what extent they provided review notes (see Appendix, p. 155). Subjects instructed to write review notes indicated a mean of 5.18; whereas, subjects who were not instructed to provide review notes indicated a mean of .93. This difference was highly significant ($t = 10.40$, $p < .001$), indicating that the manipulation was successful.

To help see if review notes proxied for counterexplanation, an exit question asked subjects to classify the review notes they listed as either opinion formation or documentation (see Appendix, p.156). According to Rich et al. (1994), opinion formation includes review notes that instruct preparers to perform “additional audit work or follow-

up,” while working paper documentation includes review notes that instruct preparers to provide “better documentation.” Subjects in the review note and combined conditions indicated that a majority of the review notes related to opinion formation (means of 75% for additional evidence and 25% for additional documentation). This result indicated that subjects believed evidence was lacking. However, independent classification results described in Section 4.2 showed that these subjects rarely were able to state explicitly what was wrong with the stated conclusion or what other explanation(s) may account for the fluctuation. Thus, it does not appear that review notes proxied for counterexplanation.

Finally, the effect of using a decision aid had a robust impact on the dependent construct both when it was used alone and in conjunction with writing review notes. However, independent of the hypotheses test results, there is no way from the data collected to ensure that the manipulation was effective other than (1) the finding that no measured covariates were driving the results and (2) the random assignment of subjects.

4.7 Analysis of Alternative Explanation Listings

The first question of Part II of the experiment, which was not given to subjects until after Part I had been collected, instructed reviewers to list all alternative explanations considered, if any (see Appendix, p. 151). For each alternative explanation listed, subjects were to state whether they believed the explanation had been ruled out. As mentioned in Section 4.1, the number of alternative explanations not ruled out did not

load onto the factor used for primary analysis and was excluded. However, further analyses were conducted on the alternative explanations listed.

While the net number of explanations not believed to have been ruled out failed to fit the hypothesized model using an ANOVA with reversed contrast weights ($F = .96$, $p < .31$), the number of total alternative explanations generated were greater when a decision aid was employed.⁴⁰ The cell means for the total number of alternative explanations listed were 1.29, 1.47, 2.17, and 1.58 for the control, review note, decision aid, and combined conditions, respectively. Although these means are in the hypothesized relationship order, they fail to fit the hypothesized model using an ANOVA with reversed contrast weights ($F = 2.57$, $p < .12$). However, combining subjects who were not provided with the decision aid (control and review note conditions) and comparing them with subjects who were provided with the decision aid (decision aid and combined conditions) leads to the finding that use of the decision aid resulted in a greater number of alternative explanations listed. A one-tailed t-test comparing the means of 1.38 and 1.87 for those subjects without and with the decision aid, respectively, was significant ($t =$

⁴⁰ Contrast weights are reversed because the variable in question is a measure of the lack of quality; whereas, other variables were a measure of quality. Also, the net number of explanations not believed to have been ruled out was measured by subtracting the number of alternative explanations believed to have been ruled out from the number of alternative explanations not believed to have been ruled out.

1.63, $p < .06$). This result suggests that the decision aid better enabled subjects to generate and perhaps consider alternative explanations. However, the absolute numbers of alternative explanations generated were quite small within all conditions.

There are several reasons why reviewers might not have listed many alternative explanations, and there is no way to tell from the data gathered which, if any, of the reasons was the most likely. First, reviewers might have been using a confirmatory processing strategy and, therefore, they might have only been looking for evidence in the working papers to support the documented conclusions.⁴¹ The possibility also exists that once they read the documented conclusions in the working papers, reviewers had difficulty considering other explanations, which is called output interference. Heiman (1990) found a similar result, which she attributed to the same psychological phenomenon, for preparers of analytical procedures. Or, the possibility exists that reviewers did not exert sufficient effort or completely understand the requirement given that they are not accustomed to listing alternative explanations when reviewing working

⁴¹ Ad hoc analysis of the review notes provided by reviewers in the review note and combined conditions (see Section 4.2) suggested that reviewers were focusing on supporting the conclusions given rather than ruling out competing, misstatement explanations.

papers.⁴² One other possibility was that the reviewers could not remember what alternatives that they had considered when reviewing the working papers. The reason that Part I of the experiment was collected prior to completing this requirement was because the act of considering and listing alternative explanations could have impacted the other dependent measures. Specifically, listing alternatives could have caused a counterexplanation effect, which would affect the other dependent measures in a manner similar to the predictions made under the Hypothesis 1b (see Chapter Two, Section 2.2.1). Thus, further research is necessary to better understand why reviewers do not generate sufficiently long lists of alternative explanations in auditing experiments such as this dissertation and Heiman (1990).

4.8 Summary of Chapter

This chapter describes how the data gathered was analyzed to test the hypotheses developed in Chapter Two. A dependent factor was created based on results of a confirmatory factor analysis used to test the covariation of the dependent measures. This factor, the summed responses to four questions representing reviewer judgment during the review process, was used to proxy the reviewer's overall judgment of the conclusions documented in the working papers. A higher overall judgment was posited to indicate

⁴² Participants in the pilot test did not mention that this requirement caused any difficulty; however, they listed a similar number of alternatives, on average, as the reviewers in the actual experiment.

that the reviewer was more likely to commit a Type II error, and a lower overall judgment was posited to indicate that the reviewer was less likely to commit a Type II error.

Classification of the review notes by independent reviewers demonstrated that instruction to provide review notes did not cue counterexplanation or generation of alternative explanations. Thus, the hypotheses tested were those that were expected to hold if the review note process were an additional cognitive task requiring capacity.

Contrast coding was used to test a model that simultaneously evaluated all five pair-wise hypotheses. The contrast was significant at the $p < .03$ level. The results suggest that the decision aid served as a retrieval plan and better enabled reviewers to counterexplain and consider alternative explanations to the documented conclusions. This argument is based on the significantly increased skepticism demonstrated by reviewers in the decision aid condition. However, the effect of the decision aid was eliminated when reviewers were also instructed to provide review notes in the combined condition. Thus, instruction to provide review notes appeared to place additional demands on reviewers' capacity to the extent that the retrieval plan failed to help increase counterexplanation or generation of alternative explanations when reviewing the working papers.

Results of the additional tests demonstrated that reviewers responded to the manipulations when they were likely to commit Type II errors *a priori*. This likelihood was measured by gathering data on reviewer perception of the client and preparer, which should have reflected low levels of reporting and knowledge bias. Also, reviewers who were more concerned with efficiency than effectiveness when reviewing were argued to

be more likely, *a priori*, to commit a Type II error. Other data gathered and analyzed failed to impact the results. However, reviewers provided the decision aid were found to generate significantly more alternative explanations than reviewers not provided the decision aid.

Chapter Five

CONCLUSION

This dissertation tested whether the audit review process for working papers can help reduce the risk of Type II errors by offsetting the decision making biases encountered when performing analytical procedures. A Type II error occurs when an auditor incorrectly accepts a client explanation indicating that no error or irregularity has occurred, when, in fact, one or the other has happened. The impact of two external cues, the instruction to write review notes and the provision of a cognitive decision aid to enhance memory retrieval, was examined between groups of audit reviewers with specialized knowledge in the financial services industry.

The results indicated that the decision aid reduced the risk of Type II errors; however, the instruction to provide review notes increased the risk of Type II errors. Using Shiffren and Shneider's (1977) information processing model and Kahneman's (1977) capacity model of attention, these affects might have occurred because of the reviewers' limited attentional capacity. Reviewers who performed the review note process might have experienced an excessive strain on capacity that increased the risk of committing Type II errors. Reviewers who were provided a decision aid might have experienced capacity relief that decreased the risk of committing Type II errors. These effects were offset when reviewers were exposed to both a cognitive decision aid and instruction to provide written review notes.

This chapter contains a summary of the dissertation study and the results of tests of the hypotheses, contributions and limitations of the study, and suggestions for future research. The chapter is divided into six sections. Section 5.1 contains a summary of the study. Section 5.2 summarizes the hypotheses and results. Section 5.3 describes contributions of the dissertation. Section 5.4 provides limitations for the experiment and results. Section 5.5 suggests future research topics. Section 5.6 concludes the dissertation.

5.1 Summary of Study

The audit review process is a key quality control function of any audit engagement and is required by generally accepted auditing standards (AICPA 1972, 1978). The audit review process consists of the evaluation of working papers that have been prepared and signed-off as complete by a subordinate auditor (preparer). The first review is performed at a detailed level by a senior or a manager, with additional reviews most likely performed at a more general level. The working paper should be defensible and justifiable if reviewed by another party, such as an engagement partner, a second partner, an external peer, or a court of law (Davis and Solomon 1989).

This dissertation examined the effectiveness of the working paper review process as a quality control mechanism to reduce the probability that a material misstatement⁴³ of

⁴³ In this study, the term “misstatement” refers only to errors or irregularities that have a material impact on the financial statements.

the financial statements is not detected during the audit. The role of the review process examined in this study was whether it was able to help auditors ensure that nonmisstatement explanations from clients were not accepted when misstatements actually might have occurred.

To help answer the research question, an audit task that relies heavily on the review process as a means of quality control was selected. The task, substantive analytical procedures, involves auditor comparisons of quantifiable data to expectations. When material fluctuations from expectations occur, the auditor must investigate the fluctuation and gather evidence to support the conclusion. Substantive analytical procedures involve the use of analytical procedures as primary sources of evidence for account balances. They are not required but provide auditors with an efficient way to gather evidence (Arens and Loebbecke 1994). Ameen and Strawser (1994) found that auditors are devoting a significant amount of time during substantive testing to analytical procedures, as high as 46% of the time for Big Six firms.

Substantive analytical procedures were important to use as a context within which to study the working paper review process because (1) they are used as a primary source of evidence for account balances and (2) they involve the auditor's ability to test the client's hypothesis for why an unexpected fluctuation occurred. The use of these procedures as primary evidences means that auditors rely on the results of the analytical procedures when concluding whether the accounts in question are stated fairly. The auditor's ability to test the client's hypothesis is relevant because should the auditor

accept a nonmisstatement explanation from the client when a misstatement actually happened, a Type II error will have occurred and the audit firm is exposed should users incur losses. These errors can prove costly to auditors, if failure to detect a misstatement results in litigation brought against the auditor by users of the financial statements (Palmrose 1991).

Failure to appropriately consider and rule out alternative hypotheses of misstatements when the client provides a nonmisstatement explanation substantially increases the risk of Type II errors. Research has demonstrated that preparers focus primarily on seeking confirmatory evidence (see Church 1990 for a review in auditing). Therefore, reviewers are needed to challenge whether preparers have gathered enough evidence to rule out misstatement explanations. Thus, a primary objective of the review process is to minimize the risk of Type II errors by evaluating the extent to which relevant alternative hypotheses have been considered and ruled out by the evidence documented in the working papers.

Because the working paper review process is an explicit quality control function, the organization of the task and differences in knowledge should better enable a reviewer than a preparer to consider alternative hypotheses for analytical procedure results. Reviewers are trained to ensure that preparers' results are consistent with the conclusions to be presented in the auditor's report (AICPA 1978). Libby and Trotman (1993) provide some evidence for this claim by demonstrating that reviewers who evaluated the information used for preparers' initial conclusions recalled more information that was

inconsistent than consistent with the preparers' conclusions. Furthermore, Frederick (1991) and Libby and Frederick (1990) demonstrate that those auditors who perform detailed working paper review, seniors and managers, have more developed knowledge than more inexperienced auditors of transaction cycles and alternative hypotheses.

However, it is possible that the number of considerations involved in the working paper review process for analytical procedures can make it difficult for reviewers to consider alternative hypotheses for cognitive reasons. In general, there is agreement in cognitive psychology that individuals have limited attentional capacity, which places a constraint on the access to stored knowledge at any given point in time (Ashcraft 1994). As part of the working paper review process for analytical procedures, a reviewer must evaluate the preparer's hypothesis, quality and extent of evidence gathered, relevance of evidence gathered, organization and preparation of the working papers, and relationship of the preparer's conclusions to other findings in the audit. In addition, reviewers must evaluate the risks associated with receiving evidence from two persuasive communicators, the client and the preparer (Rich, Solomon, and Trotman 1994). These risks involve assessments by the reviewer of the extent to which the client and preparer are competent and reporting truthfully (Rich et al. 1994, Eagly and Chaiken 1993).

It is reasonable based on extant research (cf. Libby 1985 and Libby and Frederick 1990) to assume that the experienced auditors who perform working paper reviews have the knowledge to consider alternative misstatement hypotheses. The question remains of whether the review process is affected by the difficulty that the reviewer may have

considering all aspects of the review at the same time. It is possible that reviewers are not effective at organizing internally the different considerations. Such inability can impede information processing and long-term memory access (Shiffrin and Schneider 1977).

In terms of Kennedy's (1993, 1995) debiasing framework, tests are needed to see whether the review process creates excess demands on cognitive effort or aids internal data (stored knowledge) retrieval to perform the review process for analytical procedures. Specifically, it is important to understand better the extent that reviewers counterexplain or generate and evaluate alternative hypotheses to preparers' conclusions. In this study, two external cues were introduced to help gain a better understanding of the cognitive demands placed on the reviewer when examining working papers for substantive analytical procedures.

The first cue was the instruction to provide written review notes and was included because it represents current practice in which reviewers currently write review notes to document working paper deficiencies.⁴⁴ These review notes provide preparers with explicit instructions on what follow-up procedures need to be performed and reviewers with a documented audit trail for follow-up review procedures. Instruction to provide review notes might cue counterexplanation and aid the consideration of alternative

⁴⁴ These review notes are not included in the working papers. Their sole purpose is to communicate follow-up procedures from reviewer to preparer. Due to liability concerns, these notes are kept only until the end of the audit and then destroyed.

hypotheses. Thus, the review notes themselves might proxy for counterexplanation. However, if instruction to provide review notes does not cue counterexplanation, the requirement to state explicitly any noted deficiencies and instruct the preparer on a course of action is an additional task to perform. This additional requirement would increase the amount of cognitive effort, defined as available attentional capacity or motivation (Kennedy 1995, 1993), needed for the review process. Thus, the process of considering and writing review notes might place an added strain on attentional capacity, which is limited, or reduce the motivation of the reviewer. This study controlled for changes in motivation to help identify which component of cognitive effort was affected by instruction to provide review notes.

The second cue was the provision of a cognitive decision aid containing general audit-firm guidance on how to evaluate analytical procedures. The cognitive decision aid was designed to create a retrieval plan (Medin and Ross 1992) for reviewers, which should have helped them internally organize the considerations for analytical procedures. The improved organization should have freed up limited attentional capacity to have better enabled consideration of alternative misstatement explanations. Informal, unstructured guidance like that used for this retrieval plan has been shown to enhance performance of analytical procedures by preparers (McDaniel and Kinney 1995; McDaniel 1990; and Bonner, Nelson, and Libby 1996).

To help answer whether the current review process is effective and/or can be improved, reviewers were exposed to either one, both, or neither external cue in a

between-subjects experiment. If instruction to provide review notes cued counterexplanation, then the current review process should help reduce the likelihood of committing Type II errors. Furthermore the introduction of a cognitive decision aid should have further decreased the risk of committing such errors. Conversely, if instruction to provide review notes did not cue counterexplanation, then the current review process should increase the likelihood of committing Type II errors. However, the introduction of a cognitive decision aid should help offset this difficulty by reducing the likelihood of committing Type II errors.

5.2 Summary of Hypotheses and Results

The working paper review process for analytical procedures can be described under Kennedy's (1993, 1995) debiasing framework as consisting of both effort (attention capacity and motivation) and data (internal and external). Because the review process is an explicit quality control process that can result in material economic losses to the firm if not performed appropriately, an assumption is made that reviewers are motivated sufficiently to perform the working paper review.⁴⁵ The unanswered question is whether they have sufficient capacity to perform the review, given the number of cognitive activities that reviewers perform when reviewing working papers. Furthermore, the role

⁴⁵ However, the experiment was designed to rule out motivation as a cause of any effects found between experimental conditions. Also, several measures for motivation were gathered and tested as covariates.

of the requirement to provide review notes is not known. The process of writing review notes might relieve the demands on reviewers by cueing them to counterexplain and consider alternative explanations. However, the review note process might be an additional cognitive task that increases the demands on the reviewer. The cognitive decision aid, however, was designed to relieve demands on capacity so reviewers could have more capacity available to consider alternative explanations. First, an hypothesis involving what role review notes play is discussed. Then, an hypothesis involving the decision aid is introduced. These hypotheses are followed by hypotheses that involve various comparisons and combination of the cues. Finally, the results of a contrast coding analysis that tested all five hypotheses is discussed.

Because the instruction to provide review notes is currently used in practice, this dissertation attempted to understand better its effect on reviewers; whereas, the cognitive decision aid was designed to relieve capacity demands on reviewers to enable better data access. The cognitive process of writing review notes should have reduced the risk of committing Type II errors for reviewers who do not have excessive demands on capacity from performing the review. These reviewers should have been able to write down possible working paper deficiencies, including counterexplanation of documented conclusions and the failure to rule out alternative hypotheses. According to the counterexplanation effect (Anderson et al. 1980, Koonce 1992), writing out arguments against the hypothesis should decrease the reviewer's likelihood of accepting the documented conclusions as complete. However, if there are excessive demands on a

reviewer's capacity as a result of the review process, the instruction to explicitly state working paper deficiencies should have placed even more demands on capacity. This increased strain on capacity can result in an increased likelihood of committing a Type II error because the reviewer is less likely to consider alternative explanations.

Thus, hypothesis one involved first dividing subjects according to whether their review notes contained counterexplanation or alternative hypotheses not ruled out. Reviewers who provide review notes that contain either type of information were predicted to be less likely to commit Type II errors than reviewers who did not receive either external cue. Reviewers who do not provide any counterexplanation or alternative hypotheses in their review notes were predicted to be more likely to commit Type II errors than reviewers who did not receive either external cue. However, results from an independent analysis of the review note listings showed that few reviewers provided any counterexplanation or alternative explanations. Furthermore, those review note listings that contained this information did not contain much of either type and, thus, did not have the predicted effects of the counterexplanation effect. Thus, all subjects were combined and tested under the prediction that considering and providing review notes is an additional cognitive task that requires effort and thus capacity.

The cognitive decision aid was designed to create a retrieval plan, which is an organized set of cues to aid the retrieval of information (Medin and Ross 1992). Because the decision aid provided a sequential means for considering working papers in order to prevent Type II errors, the second hypothesis predicted that reviewers provided with the

decision aid would be less likely to commit Type II errors than reviewers who only were instructed to provide review notes. This result would suggest that the decision aid works as intended: it enabled the reviewer to internally organize the considerations for reviewing analytical procedures, which freed up capacity to consider alternative explanations. However, the possibility exists that the provision of a decision aid caused the reviewer to increase his or her motivation to consider alternative explanations. Hypothesis five, discussed below, was designed to rule out this competing explanation.

Hypothesis three involved a comparison between the impact of provision of the decision aid to the instruction to write review notes. The hypothesis predicted that reviewers who only were provided the decision aid would be less likely to commit Type II errors than reviewers who only were instructed to write review notes. Because instruction to provide review notes was predicted to cause an increase in likelihood to commit Type II errors, hypothesis three should have been supported because each cue should have opposite impacts on the likelihood to commit Type II errors.

Hypotheses four and five involved predictions for reviewers who were provided with both external cues. These hypotheses were needed to demonstrate the robustness of findings for hypotheses one and two. Also, these hypotheses help rule out competing explanations for results of the first two hypotheses.

Hypothesis four examined the impact of adding a decision aid to the current practice of instructing reviewers to provide review notes. The hypothesis predicted that the provision of a decision aid would decrease the likelihood of committing a Type II

error when compared to reviewers who only have been instructed to provide review notes. Because review notes were predicted to increase the likelihood of committing Type II errors, the provision of a decision aid should offset, or at least reduce, the negative effects of review notes.

Hypothesis five was perhaps most important because it served as a strong indicator of the role that review notes plays in the review process. Because review notes were predicted to increase the likelihood of committing Type II errors for hypothesis one, reviewers who receive both external cues were predicted to be more likely to commit Type II errors than reviewers who only are provided a decision aid. This prediction was based on the argument that the instruction to provide review notes required the reviewer to perform an additional, difficult cognitive task. If the decision aid was found to decrease the likelihood to commit Type II errors in hypothesis two, then the additional requirement to provide review notes should not have increased the likelihood to commit Type II errors, unless providing review notes was an additional cognitive task. In other words, the reviewer had sufficient capacity to provide written review notes because of the decision aid. Therefore, the requirement to provide review notes should only increase the likelihood of Type II errors if that requirement places an additional strain on capacity such that part, or all, of the benefit of the decision aid was eliminated.

Thus, an increase in the likelihood to commit Type II errors for hypothesis five, in conjunction with results from hypothesis one, would provide strong evidence that the instruction to provide written review notes is a difficult cognitive task. As a result, the

possibility of Type II errors resulting from the preparation and review of substantive analytical procedures would be a legitimate concern that auditing firms need to investigate further.

To test the hypotheses, a between-subjects experiment was conducted in which 78 seniors and managers from two international accounting firms reviewed working papers for a substantive analytical procedure from the banking industry. A nonmisstatement explanation for the material fluctuation from prior years and the industry was provided by the client and documented by the preparer. The evidence gathered by the preparer to support the explanation was not sufficient to rule out competing misstatement explanations or support the given explanation. Thus, reviewers should have been critical of the working papers.

The dependent measure for the hypotheses tests, which was selected based on the results of a confirmatory factor analysis, was the sum of four questions relating to the validity, persuasiveness, sufficiency, and type of evidence gathered by the preparer to support the documented explanation. These questions loaded onto one dependent factor, which was posited to represent the reviewer's overall judgment of the conclusions documented in the working papers. A high overall judgment was argued to be indicative of a higher likelihood to accept the working papers and, thus, a higher risk of committing a Type II error if the "true" explanation involved a misstatement. A low overall judgment was argued to be indicative of a lower likelihood to accept the working papers and, thus, a lower risk of committing a Type II error.

Contrast coding was used to test a model that simultaneously evaluated all five pair-wise hypotheses by assigning contrast weights of 0, 1, -1, and 0 to cell means of the dependent factor for the control, review note, decision aid, and combined conditions, respectively.⁴⁶ The contrast was significant at the $p < .03$ level. The results suggest that the decision aid served as a retrieval plan and better enabled reviewers to counterexplain and consider alternative explanations to the documented conclusions. This argument is based on the significantly increased skepticism demonstrated by reviewers in the decision aid condition. However, the effect of the decision aid was eliminated when reviewers were also instructed to provide review notes in the combined condition. Thus, instruction to provide review notes appeared to place additional demands on reviewers' capacity to the extent that the retrieval plan failed to help increase counterexplanation or generation of alternative explanations when reviewing the working papers.

Results of the additional tests demonstrated that reviewers responded to the manipulations when they were likely to commit Type II errors *a priori*. This likelihood was measured by gathering data on reviewer perception of the client and preparer, which

⁴⁶ The experimental conditions consisted of the following: the control condition contained neither experimental cue, the review note condition contained only the instruction to provide review notes, the decision aid condition contained only the provision of a decision aid, and the combined condition contained both external cues.

should have reflected low levels of reporting and knowledge bias.⁴⁷ Also, reviewers who were more concerned with efficiency than effectiveness when reviewing were argued to be more likely, *a priori*, to commit a Type II error. Other data gathered and analyzed failed to impact the results. However, reviewers provided the decision aid were found to generate significantly more alternative explanations than reviewers not provided the decision aid.

5.3 Contributions

This study makes several contributions to the study of decision making in auditing. The dissertation evaluates whether the review process for analytical procedures helps to reduce the risk of Type II errors. The review note process provides the opportunity to examine whether the current review setting places a strain on available capacity to consider alternative hypotheses or counterexplanation. The use of a cognitive retrieval plan is included to see if it aids internal data access and eases available capacity

⁴⁷ Reporting bias and knowledge bias are discussed in detail in Chapter Two, Section 2.3.2. As a reminder, reporting bias refers to the degree to which a communicator's willingness to convey an accurate message is compromised, and knowledge bias refers to the degree to which a communicator's knowledge of message relevant information is inaccurate (Eagly and Chaiken 1995, 1984). The background information on the client and preparer was designed so that reviewers would evaluate both reporting and knowledge bias as low.

strain. By comparing and combining both cues and including a control condition, it is possible to obtain an understanding of the extent to which reviewers are able to access data, given possible strains on capacity.

The results suggest that reviewers who are likely to commit Type II errors and are instructed to provide review notes may have difficulty counterexplaining and considering alternative hypotheses without being cued with external guidance. Reviewers who provided review notes rated the experimental working paper conclusions as being of higher overall quality than those who had no cue. However, the addition of a cognitive decision aid offset the difference.

The difficulty reviewers had counterexplaining and retrieving alternative hypotheses when providing review notes can be seen further when looking at the comparison of those who only read the decision aid to those who were exposed to both manipulations. Overall, subjects who only read the decision aid were more critical of the working papers than those reviewers who also provided review notes.

Perhaps the inclusion of a cognitive decision aid would be a useful addition to the review process. As working papers become more automated, such an aid can be included as a pop-up screen on a computer prior to the review of on-line working papers containing analytical procedures. Since the decision aid only takes a few minutes to study, it would be relatively inexpensive to implement into the standard review process for analytical procedures.

The measurement of the dependent construct using multiple variables to measure better the dependent construct is also a contribution. Past studies have used a single measure when assessing the judgments that auditors make when evaluating an explanations that has been provided for an analytical procedure. This study identified three other variables that can be measured along with the validity measure used in past studies (Heiman 1990, Koonce, 1992, Peecher [forthcoming], and Phillips 1996) to proxy for the auditor judgment construct. Use of confirmatory factor analysis to test this multiple measure construct demonstrated that measurement error of the construct can be reduced by employing these four measures instead of only the one previously used.

Another important contribution is the demonstration of the importance of controlling for environmental factors surrounding *a priori* likelihood of committing Type II errors when performing studies on the working paper review process. Reviewer perceptions of the client and preparer concerning their knowledge and/or propensity to lie appear to affect the strategy undertaken by the reviewer. High concerns for these issues might cause reviewers to focus harder on the quality of the evidence to support the given explanation or rule out competing explanations. However, low concerns for these issues might cause reviewers to focus harder on documentation or other issues surrounding the working papers. While this might be an appropriate strategy in most circumstances, the *risk* of a Type II error actually increases in the latter scenario because the occurrence of a material misstatement is a rare event. Thus, although the probability of a misstatement may be less for clients with a low level of reporting or knowledge bias, the probability of

detecting such a misstatement might be much less without the help of cognitive guidance. Conversely, though the probability of a material misstatement is higher for a client with a high level of reporting or knowledge bias, the probability of *detecting* the misstatement should be much greater given that reviewers are more focused on detecting the misstatement.

The dissertation also extends Heiman (1990) to see how well reviewers generate alternative hypotheses upon inheriting hypotheses from preparers. Use of the decision aid, which included instructions to consider alternative hypotheses, significantly increased the number of alternative hypotheses listed by reviewers. This comparison of the decision aid with the instruction to list review notes also allows for a comparison of the effects of self-generated alternative hypotheses (written review notes in this study) and those that are generated based on experimenter-provided cues of general causation of fluctuations (provision of decision in this study). Such a comparison was suggested in Heiman (1990) as a topic for future research; however, she was specifically discussing the use of an actual list of hypothesized causes as opposed to a decision aid designed to help facilitate the retrieval of hypothesized causes.

5.4 Limitations

There are several limitations to this study. First, the actual review process in this study does not involve any discussion with the preparer. Prior research (Ismail and Trotman 1994) has shown that the review process is more accurate when discussion between preparer and reviewer is allowed. Because this study is concerned with the

impact of external cues on the reviewer's cognitive processes, it is important to hold constant the information available to the reviewer.

Second, the review process is being examined only within the context of substantive analytical procedures. The review process for other audit areas is likely to be affected differently by external cues and modeled in a different fashion (e.g., whether the account balance was confirmed sufficiently would be considered more important for accounts receivable).

Finally, as with many research tasks of this sort, generalizations from the results of this study are limited because of the controlled environment in which it is administered and the limited number of subjects involved. Reviewers have no other working papers to examine, and they do not actually know the preparer or the client. The absence of such factors (that exist in a natural setting) could have an impact on the interpretation of results of this study.

5.5 Suggestions for Future Research

The results found in this dissertation provide several research questions that can be addressed in future studies. This section briefly discusses these suggestions and discusses possible implications for the results contained in this dissertation.

One important question involving the working paper review process is the cognitive demands of performing the review, absent any external cues. Representatives at the firms involved in this dissertation have suggested that the review process is likely to change. One possible change is the elimination of the *ex post* traditional review note

process. Instead, reviews will occur concurrent with the preparation process, so that the conclusion by the preparer will already be approved by the reviewers. This change is concurrent with an overall streamlining of the documentation contained in working papers: a recommendation firms are implementing based on advice from legal counsel. Less documentation in working papers means that reviewers need to be more involved with preparers during the preparation of the working papers.

The results in this dissertation were based on comparisons to a control condition for Hypotheses 1 and 2. Results appeared to suggest that subjects in the control condition were experiencing a strain on capacity (i.e., they were not critical of the insufficiently supported conclusions documented in the working papers. Thus, future research is needed that examines the cognitive processes involved in the working paper review process.

Specifically, the results in this dissertation suggest that reviewers in this task, which closely modeled a working paper review in the natural environment when the preparer is not present, might have experienced excessive demands on their capacity. Perhaps a more abstract working paper review task that gathers direct evidence relating to capacity overload would provide additional support for the arguments posited in this dissertation.

Also, a within-subjects experiment might help test whether the decision aid provided a retrieval plan that eased demands on capacity and better enabled counterexplanation and generation of alternative explanations. While the between-

subjects experiment helps eliminate the increased possibility of demand effects associated with within-subject experiments, it is difficult to determine whether subjects were aware of the effects of the decision aid that occurred.

Because of limitations on subject availability, there were several factors held constant that would have been interesting research questions had the variables been manipulated. First, the reporting and knowledge bias variables could have been varied to see if (1) reviewers would have responded to the manipulations for each type of bias and (2) reviewers would have responded differently to the experimental cues as predicted. Second, testing reviewer behavior for a different type of auditing test could help identify the reviewer's primary concerns for that task, which should be focused more on the sufficiency of evidence to support, for example, a material portion of an asset balance. Also, allowing discussion with a preparer and examining the effects of group interaction on reviewer behavior would provide insights to whether reviewers also exhibit excessive strains on capacity when discussion is allowed.

5.6 Concluding Statement

This dissertation represents an initial look at the impact of the instruction to provide review notes on the working paper review process, and it provides initial evidence on the impact of adding a cognitive decision aid to the review process. Thus, the task was designed to be as close to a natural task as possible to understand whether these issues warrant future study. This dissertation also is one of the first auditing studies to address the decision-making process from the perspective of demands on capacity. These demands occur at the initial information processing stage, prior to retrieval of

information from long-term memory. Accordingly, this experiment introduces as many research questions as it helps to answer. All of the findings in this dissertation should be further supported with more abstract tasks that more directly test the hypothesized effects of the theoretical constructs for this aspect of the working paper review process. Only then can the two studies be combined to have a clearer understanding of the working paper review process for substantive analytical procedures.

APPENDIX

Appendix

PART II OF EXPERIMENTAL INSTRUMENT: ADDITIONAL DATA

The following appendix is a copy of Part II of the experimental instrument used for the dissertation. The questions are identical to what the reviewers actually were given during the experiment. Only the margins and scale sizes have been changed to meet established, dissertation guidelines.

ADDITIONAL QUESTIONS RELATED TO CASE

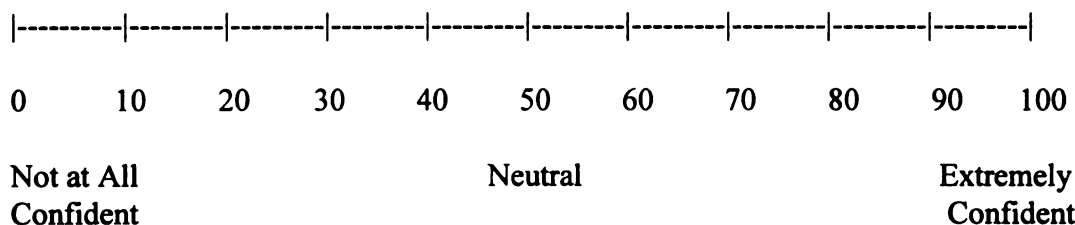
1. When evaluating the working papers, what, if any, alternative explanations (based on your training and experience) did you consider for the decrease in gross interest income as a percentage of average earning assets? To the extent you considered these alternative explanations, did you believe the corroborative evidence gathered ruled them out? **Please number and document any alternative explanations considered and indicate whether you believed them to be ruled out by the evidence gathered.**

Alternative Explanation

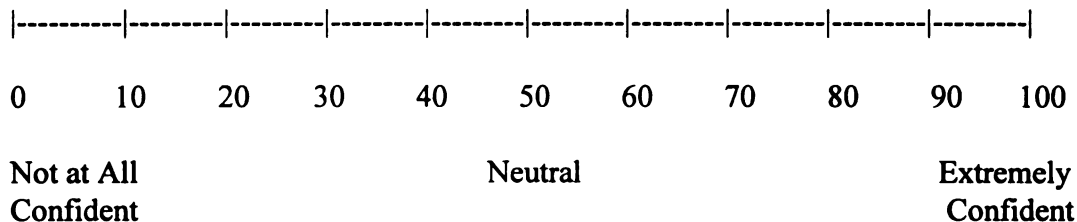
Ruled out by
Corroborative
Evidence?
(Yes/No/Unsure)

1.

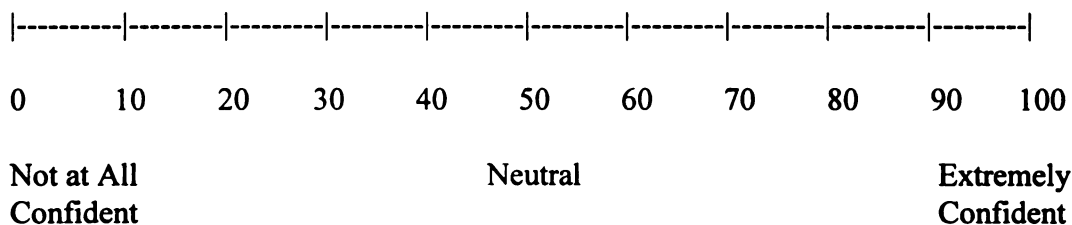
2. How confident are you about *your estimate* of the probability that the documented explanation accounts for substantially all (e.g., ≥ 85 percent) of the decrease in gross interest income as a percentage of average earning assets?



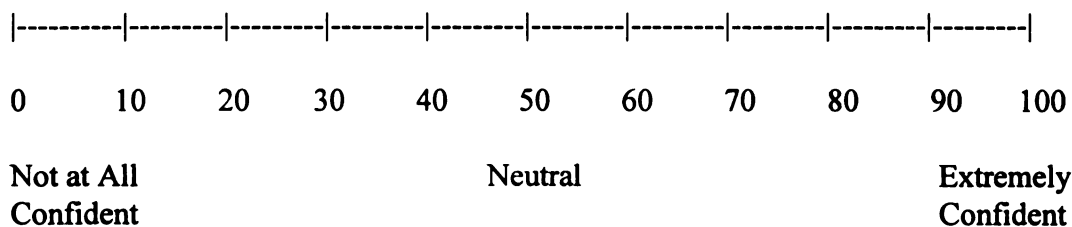
3. How confident are you about *your estimate* of the extent to which additional corroborative evidence needs to be gathered before you are willing to sign off?



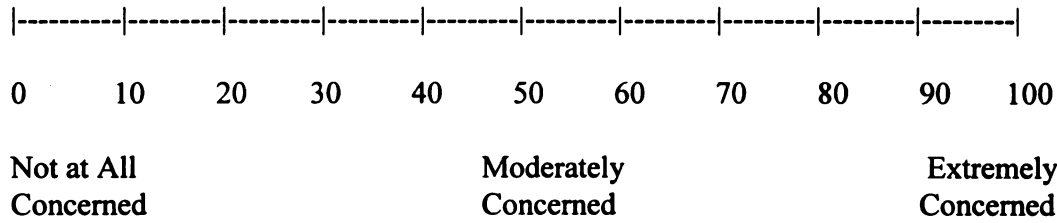
4. How confident are you about *your estimate* of the *extent* of corroborative evidence previously gathered by the preparer?



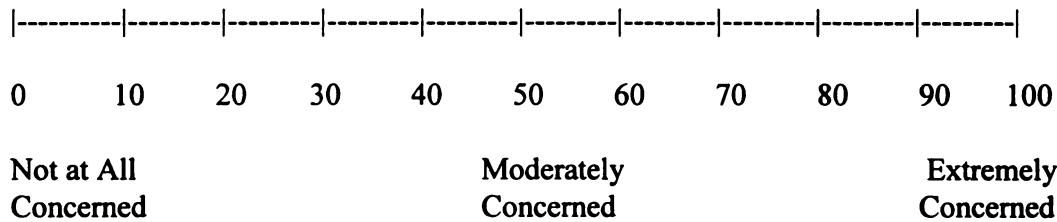
5. How confident are you about *your estimate* of the *nature* of corroborative evidence previously gathered by the preparer?



6. When reviewing the analytical procedure for gross interest income as a percentage of average earning assets, how concerned were you about accepting client explanations without adequate justification?

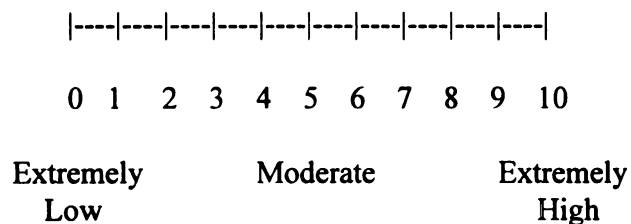


7. Considering your answer to Number 6, how concerned were you about performing costly investigations without adequate justification?



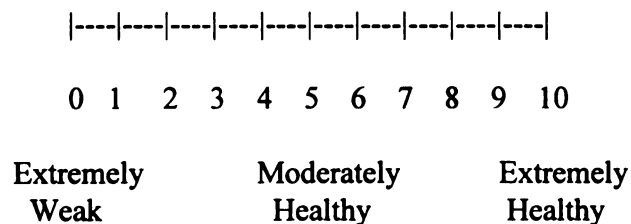
8. Take a few moments to reflect on the background information of the case. Now, evaluate the following factors as you *perceived* them when reviewing the working papers.

a. The audit risk associated with the client.


☐

Check here if you had no perception of this issue when performing the review.

b. The financial condition of the client.


☐

Check here if you had no of this issue when perception performing the review.

- c. The effectiveness of the internal control structure.

--- --- --- --- --- --- --- --- ---										
0	1	2	3	4	5	6	7	8	9	10
Extremely Ineffective			Moderately Effective				Extremely Effective			

☐

Check here if you had no perception of this issue when performing the review.

- d. The competence of the client.

--- --- --- --- --- --- --- --- ---										
0	1	2	3	4	5	6	7	8	9	10
Extremely Low			Moderate				Extremely High			

☐

Check here if you had no perception of this issue when performing the review.

- e. The integrity of the client.

--- --- --- --- --- --- --- --- ---										
0	1	2	3	4	5	6	7	8	9	10
Extremely Low			Moderate				Extremely High			

☐

Check here if you had no perception of this issue when performing the review.

- f. The materiality of the decrease in gross interest income as a percentage of average earning assets.

--- --- --- --- --- --- --- --- ---										
0	1	2	3	4	5	6	7	8	9	10
Not at All Material			Moderately Material				Highly Material			

☐

Check here if you had no perception of this issue when performing the review.

- g. The amount of time remaining to complete testing in this area relative to the scheduled closing meeting and earnings announcement date.

--- --- --- --- --- --- --- --- --- ---											☐
0	1	2	3	4	5	6	7	8	9	10	Check here if you had no perception of this issue when performing the review.
A Great Deal of Time Left			Moderate Time Left			Very Little Time Left					

- h. The competence of the staff auditor who prepared the working papers.

--- --- --- --- --- --- --- --- --- ---											☐
0	1	2	3	4	5	6	7	8	9	10	Check here if you had no perception of this issue when performing the review.
Extremely Low			Moderate			Extremely High					

- i. The amount of time used by the staff accountant relative to total time in the budget for the test.

--- --- --- --- --- --- --- --- --- ---											☐
0	1	2	3	4	5	6	7	8	9	10	Check here if you had no perception of this issue when performing the review.
Insufficient Time Spent			Just Enough Time Spent			Too Much Time Spent					

9. Prior to providing answers for the questions in which you evaluated the conclusions in the working papers (the questions in Part I of the case that were previously turned in), to what extent did you explicitly write review notes?

----- ----- ----- ----- ----- ----- ----- ----- ----- -----										
0	1	2	3	4	5	6	7	8	9	10
I Did Not Explicitly Write Any Review Notes			I Explicitly Wrote Some Review Notes				I Explicitly Wrote Many Review Notes			

10. Review notes are generally written either (1) as a means of improving documentation for future users of the working papers or (2) to instruct the preparer to gather additional corroborative evidence to support the explanation provided or rule out other possible explanations. If you explicitly wrote review notes, how would you classify them? (percentages should sum to 100 %)

a. Better Documentation: _____ % ☐

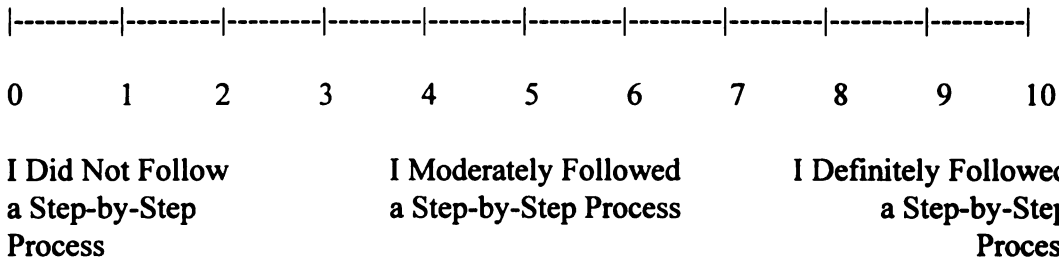
b. Additional Evidence: _____ %

100 %

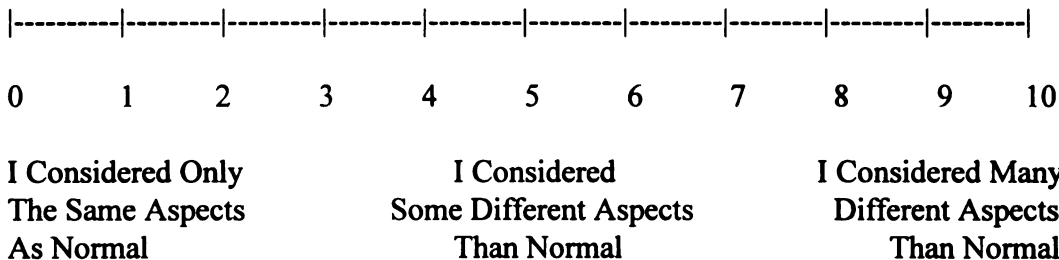
Check here if you did not explicitly write any review notes (question is not applicable).

11. Prior to providing answers for the questions in which you evaluated the conclusions in the working papers (the questions in Part I of the case that were previously turned in), to what extent did you:

a. perform a general, step-by-step process for performing a review of analytical procedures?



b. consider a particular aspect of the review process that you normally would not have considered?



c. consider alternative explanations for the fluctuation not mentioned in the working papers?

----- ----- ----- ----- ----- ----- ----- ----- ----- -----										
0	1	2	3	4	5	6	7	8	9	10
I Considered No Alternative Explanations				I Considered Some Alternative Explanations				I Considered Many Alternative Explanations		

d. consider the quantitative support for the explanation of the fluctuation documented in the working papers?

----- ----- ----- ----- ----- ----- ----- ----- ----- -----										
0	1	2	3	4	5	6	7	8	9	10
I Did Not Consider The Quantitative Support For The Explanation				I Moderately Considered The Quantitative Support For The Explanation				I Highly Considered The Quantitative Support For The Explanation		

e. consider that both the client and the preparer introduce different risks that need to be assessed when reviewing the working papers?

----- ----- ----- ----- ----- ----- ----- ----- ----- -----										
0	1	2	3	4	5	6	7	8	9	10
I Did Not Consider These Risks				I Moderately Considered These Risks				I Highly Considered These Risks		

12. Please specify any other special considerations you gave to the review of the working papers in Part I of the case:

ADDITIONAL QUESTIONS RELATED TO YOUR PROFESSIONAL EXPERIENCE

1. Sex: _____ (M or F)

2. Age: _____

Please answer the following questions using your best estimate.

3. Experience with the firm: _____ years and _____ months

4. Position in the firm during the majority of the past 12 months:

5. Experience at the position stated in Number 4: _____ years and _____ months

6. Average number of clients assigned per year while in the position stated in Number 4: _____ clients

7. Percentage of chargeable time spent reviewing working papers in the position stated in Number 4: _____%

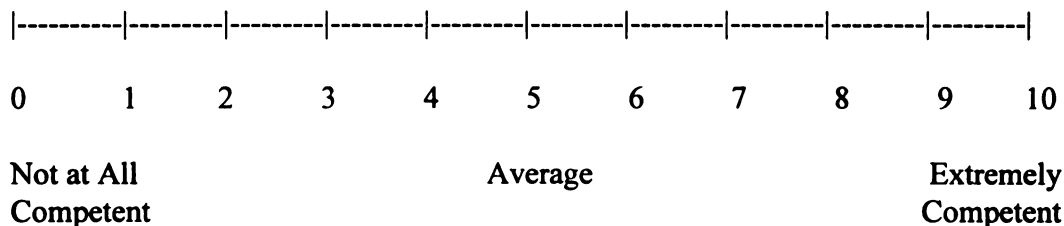
8. Percentage of chargeable time spent performing analytical procedures in the position stated in Number 4: _____%

9. Percentage of chargeable time spent reviewing analytical procedures in the position stated in Number 4: _____%

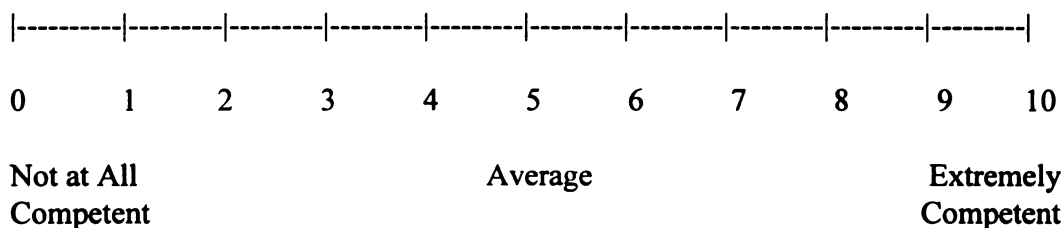
10. Percentage of chargeable time spent on clients in the financial services industry in the position stated in Number 4: _____%

11. Prior experience as an employee for a company in the financial services industry: _____ years and _____ months

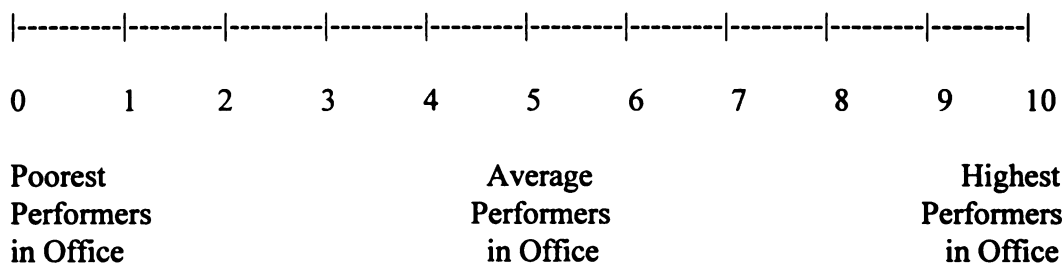
12. Overall, how competent would you rate *other reviewers* at the firm in the same position as stated in Number 4 in the ability to review analytical procedure working papers in the financial services industry?



13. How competent, relative to other reviewers at the firm in the same position as stated in Number 4, would you rate *yourself* in the ability to review analytical procedure working papers in the financial services industry?



14. On average, what is the competence level of the staff who typically prepare the working papers you review?



15. How much effort did you exert when reviewing the analytical procedure working papers?

----- ----- ----- ----- ----- ----- ----- ----- ----- -----										
0	1	2	3	4	5	6	7	8	9	10
Much Less Effort Than on an Actual Review				As Much Effort as on an Actual Review			Much More Effort Than on an Actual Review			

LIST OF REFERENCES

LIST OF REFERENCES

- Ameen, E. C. and J. R. Strawser. 1994. Investigating the use of analytical procedures: An update and extension. *Auditing: A Journal of Practice and Theory* 13 (Fall): 69-76.
- American Institute of Certified Public Accountants. 1972. Codification of Auditing Standards and Procedures: Statement of Auditing Standards No. 1. AICPA: New York.
- American Institute of Certified Public Accountants. 1978. Planning and Supervision: Statements on Auditing Standards No. 22. AICPA: New York.
- American Institute of Certified Public Accountants. 1988. Analytical Procedures: Statements on Auditing Standards No. 56. AICPA: New York.
- Anderson C., M. Lepper, and L. Ross. 1980. Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology* 39 (December): 1037-49.
- Anderson, U., L. Koonce, and G. Marchant. 1994. The audit review process and its effect on auditors' assessments of evidence from management. *Advances in Accounting* (Forthcoming).
- Arens, A. and J. Loebbecke. 1994. *Auditing: An Integrated Approach* Englewood Cliffs, N.J.: Prentice Hall.
- Asare, S. K. and L. S. McDaniel. 1996. The effect of familiarity with the preparer and task complexity on the effectiveness of the audit review process. *The Accounting Review* 71 (April): 139-59.
- Ashcraft, M.H. 1994. *Human Memory and Cognition*. HarperCollins College Publishers: New York, NY. pp 236-247.
- Ashton, A. H. 1991. Experience and error frequency knowledge as potential determinants of auditor expertise. *The Accounting Review* 66(April): 218-239.
- Bamber, L., M. Bamber, and J. Bylinski. 1988. A descriptive study of audit managers' working paper review. *Auditing: A Journal of Practice and Theory* 7 (Spring): 137-49.

- Bedard, J. C. and S. F. Biggs. 1991a. The effect of domain-specific experience on evaluation of management representations in analytical procedures. *Auditing, A Journal of Practice and Theory*. 10 (Supplement): 77-90.
- Bedard, J. C. and S. F. Biggs. 1991b. Pattern recognition, hypothesis generation, and auditor performance in an analytical task. *The Accounting Review*. 66 (July): 622-642.
- Birnberg, J. and M. Shields. 1984. The role of attention and memory in accounting decisions. *Accounting, Organizations, and Society* 9 (3/4): 365-82.
- Bloucher, E. and G. Patterson. 1995. The use of analytical procedures. *Journal of Accountancy* (February): 53-5.
- Bonner, S. and B. Lewis. 1990. Determinants of auditor expertise. *Journal of Accounting Research* 28 (Supplement): 1-20.
- Bonner, S., R. Libby, and M. W. Nelson. 1996. Using decision aids to improve auditors' conditional probability judgments. *The Accounting Review* 71 (April): 221-40.
- Bonner, S. and N. Pennington. 1991. Cognitive processes and knowledge as determinants of auditor expertise. *Journal of Accounting Literature*. 10: 1-50.
- Buckless, F. and S. Ravenscroft. 1990. Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review* 65 (October): 933-45.
- Church, B. 1990. Auditors' use of confirmatory processes. *Journal of Accounting Literature* 9: 81-112.
- Davis, J. and I. Solomon. 1989. Experience, expertise, and expert-performance research in public accounting. *Journal of Accounting Literature*. 8: 150-164.
- Eagly, A. and S. Chaiken. 1984. Cognitive theories of persuasion. In L. Berkowitz (ed.) *Advances in Experimental Social Psychology* pp. 267-359. Orlando: Academic Press.
- Eagly, A. and S. Chaiken. 1993. *The Psychology of Attitudes*. Fort Worth: Harcourt Brace Jovanovich.
- Edwards, W. 1983. Human cognitive capabilities, representativeness, and ground rules for research. In P. Humphries, O. Svenson, and A. Vari (eds.) *Advances in Psychology: Analyzing and Aiding Decision Processes* New York: North-Holland.

- Fischhoff, B., P. Slovic, and S. Lichtenstein. 1978. Fault Trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance* (May): 193-242.
- Frederick, D. 1991. Auditors' representation and retrieval of internal control knowledge. *The Accounting Review* 66 (April): 240-58.
- Frederick D., V. Heiman-Hoffman, and R. Libby. (1994). The structure of auditors' knowledge of financial statement errors. *Auditing: A Journal of Practice and Theory* 13: 1-21..
- Gibbins, M. and C. Emby. 1985. Evidence on the nature of professional judgment in public accounting. *Audit Research Symposium 1984* (Office of Accounting Research, College of Commerce and Business Administration: Champaign, IL): 181-212.
- Gillet, P. R., E. M. Bamber, T. J. Mock, and K. T. Trotman. 1994. Audit Judgment. Working Paper.
- Heiman, V. 1990. Auditors' assessments of the likelihood of error explanations in analytical review. *The Accounting Review*. 65: 875-890.
- Hogarth, R. and H. Einhorn. 1992. Order effects in belief updating: The belief adjustment model. *Cognitive Psychology* 24: 1-55.
- Ismail, Z. and K. Trotman. 1994. The impact of the review process in hypothesis generation tasks. working paper. University of New South Wales.
- Kahneman, D. 1973. *Attention and Effort* Englewood Cliffs, N.J.: Prentice-Hall.
- Kennedy, J. 1995. Debiasing the curse of knowledge in audit judgment. *The Accounting Review* 70 (April): 249-73.
- Kennedy, J. 1993. Debiasing audit judgment with accountability: A framework and experimental results. *Journal of Accounting Research* 31(Autumn): 231-45.
- Keppel, G. 1982. *Experimental Design* Blemont, CA: Brooks/Cole.
- Kim, J. and C. Mueller. 1978a. *Factor Analysis: Statistical Methods and Practical Issues* Newbury Park, CA: Sage Publications Inc.
- Kim, J. and C. Mueller. 1978b. *Introduction to Factor Analysis: What It Is and How To Do It* Newbury Park, CA: Sage Publications Inc.

- Klayman, J. and Y. Ha. 1987. Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review* 94 (2): 211-28.
- Koonce, L. 1993. A cognitive characterization of audit analytical review. *Auditing: A Journal of Practice and Theory* 12 (Supplement): 57-76.
- Koonce, L. 1992. Explanation and Counterexplanation during audit analytical review. *The Accounting Review*. 67: 59-76.
- Koonce, L., U. Anderson, and G. Marchant. 1994. Justification of decisions in auditing. *Journal of Accounting Research* (Forthcoming).
- Koriat, A., S. Lichtenstein, and B. Fischhoff. 1980. Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory* 6 (March): 107-118.
- Libby, R. 1985. Availability and the generation of hypotheses in analytical review. *Journal of Accounting Research* 23(Autumn): 648-667.
- Libby, R. and D. M. Frederick. 1990. Experience and the ability to explain audit findings. *Journal of Accounting Research* 28 (Autumn): 348-367.
- Libby, R. and J. Luft. 1993. Determinants of judgment performance in accounting settings: Ability knowledge, motivation, and environment. *Accounting, Organization, and Society* (July): 425-50.
- Libby R. and K. Trotman. 1993. Audit Review as a control for biased recall of evidence in decision making. *Accounting, Organizations, and Society* 18: 559-574.
- Macintosh, N. and J. Williams. 1992. Managerial roles and budgeting behavior. *Behavioral Research in Accounting* 4: 23-48.
- McDaniel, L. S. 1990. The effects of time pressure and audit program structure on audit performance. *Journal of Accounting Research* (Autumn): 267-85.
- McDaniel, L. S. and W. R. Kinney, Jr. 1995. Expectation-formation guidance in the auditor's review of interim financial information. *Journal of Accounting Research* 33 (Spring): 59-76.
- Medin, D. and B. Ross. 1990. *Cognitive Psychology*. Ft. Worth: Harcourt, Brace, Jovanovich.
- Mehle, T. 1982. Hypothesis generation in an automobile malfunction inference task. *Acta Psychologica* 52: 87-106.

- Mehle, T., C. F. Gettys, C. Manning, S. Baca, and S. Fisher. 1981. The availability explanation of excessive plausibility assessments. *Acta Psychologica* 49: 127-140.
- Messier, W. F. Jr. and R. M. Tubbs. 1994. Recency effects in belief revision: The impact of audit experience and the review process. *Auditing: A Journal of Practice and Theory* 13(Spring): 57-72.
- Moeckel, C. 1990. The effect of experience on auditors' memory errors. *Journal of Accounting Research* 28 (Spring): 368-87.
- Moeckel, C. and R. Plumlee. 1989. Auditors' confidence in recognition of audit evidence. *The Accounting Review* 64 (October): 653-66.
- Nelson, M., R. Libby, and S. Bonner. 1995. Knowledge structure and the estimation of conditional probabilities in audit planning. *The Accounting Review* 70 (January): 27-47.
- Palmrose, Z. 1991. Trials of legal disputes involving independent auditors: Some empirical evidence. *Journal of Accounting Research* 29(Supplement): 149-185.
- Peecher, M. E. Forthcoming. The influence of auditors' justification processes on their decisions: A cognitive model and experimental evidence. *Journal of Accounting Research*.
- Phillips, F. 1996. Evaluating financial reporting issues: The influence of audit workpaper review strategies. Working Paper. The University of Texas
- Pincus, K. 1989. The efficacy of a red flags questionnaire for assessing the possibility of fraud. *Accounting, Organizations, and Society* 14: 153-63.
- Ramsey, R. J. 1994. Senior/manager differences in audit workpaper review performance. *Journal of Accounting Research* 32(Spring) 127-135.
- Rich, J. 1996. Elaboration in audit financial-statement review: An experimental investigation from a persuasion perspective. Working Paper.
- Rich, J.S., I. Solomon, and K. T. Trotman. 1994. The audit review process: A cognitive characterization from the persuasion perspective. working paper. The University of Illinois.
- Shiffrin, R. M. and W. Schneider. 1977. Controlled and automatic human information processing: 2. Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84:127-90.

- Trotman, K. 1985. The review process and the accuracy of auditor judgments. *Journal of Accounting Research*. 23: 740-752.
- Trotman, K. and P. Yetton. 1985. The effect of the review process on auditor judgments. *Journal of Accounting Research*. 23: 256-267.

MICHIGAN STATE UNIV. LIBRARIES



31293015671898