Closing Review Notes: The Effect of Reviewer Delay and Review Note Frame on Audit Workpaper Preparers’ Effort and Performance

A Thesis
Submitted to the Faculty
of
Drexel University
by
Tamara A. Lambert
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
August 2009
DEDICATION

To Neal, my loving and patient husband who has worked tirelessly during the last five years so that I could follow this career path and then spent his days off working on the house, cooking, ironing, and providing technical and emotional support for me.
I was very fortunate to have Christopher P. Agoglia as my advisor, dissertation chair, and friend throughout my time as a doctoral student. Chris, you are an extraordinarily effective mentor and I thank you for everything you have taught me over the last several years. Your dedication to our profession and to the education of researchers is astounding and I hope to be able to emulate it throughout my career. I am grateful for all of the times that you refused to accept “good enough” and I promise to never write the word “impacted” again unless I am describing a dental condition (I cannot promise I will always remember the “with” after “done” when speaking, but I will try my best).

I am also grateful to the rest of my committee members, Anna Cianci, Nandini Chandar, John Schaubroeck, George Tsakumis, and Arnold Wright for providing their precious time, invaluable suggestions, and much-needed support throughout this process. George and Anna also provided a much-needed outlet to discuss celebrity gossip and other pop-culture phenomenon. Nandini’s encouragement and ability to make difficult concepts understandable helped me to actually enjoy her capital markets seminar (and come out of it with a viable working paper). John showed a great deal of patience as I struggled through my first exposure to academic research during the behavioral science seminar that he taught. Although he has never been formally associated with Drexel University, Arnie was kind enough to agree to be on my committee and has provided very detailed and helpful comments and suggestions. I would also like to acknowledge David Campbell and Paul Jensen for believing in me and ensuring that I received support.
for research projects and conferences. Also, Barbara Grein allowed me to use her office when she wasn’t and was a sounding board and friend throughout my time at Drexel.

My time at Drexel did not overlap with former doctoral students Joseph Brazel or Maria Sanchez. However, that has not stopped them from providing valuable help and guidance to me in navigating the academic world. They are both excellent colleagues who have made a positive impact on my career thus far. Thankfully, my time did overlap with Cathy Beaudoin, and I would have been lost during my first year without her. Other Drexel cohorts of mine, Xiaochuan Zheng, Yongtao Hong, Ross Roberts, and Robert Lee have also provided assistance and friendship. In addition, several non-Drexel doctoral students have acted as a support system for me throughout the latter half of my time in the program, Linda Kolbasovsky, Ben Luippold, and Chad Stefaniak.

This research would not have been possible without the help of audit professionals who served as advisors in developing the instrument, pre-tested various aspects of the instrument, and participated in the experiment. I am very grateful to all of them.

Finally, I would like to thank my family. In addition to my wonderful husband (to whom this is dedicated), I must acknowledge the love and support of my mother, MaryAnn Grigas. She instilled in me a love of learning and the confidence that I could do whatever I set my mind to. Also, my sister and brother-in-law, Lisa and Eamonn McDowell, and my uncle and aunt, John and Sherry Baldino, have helped me to keep my sanity throughout this process. We may be far apart in miles (3,000 or so, in opposite directions) but we are as close as any family can be. Lastly, I must mention my grandfather, Anthony Baldino, who passed away this year. He was truly like a father to
me and lived with us throughout most of my time as a doctoral student. I am grateful to have had his love and support throughout this time.
# TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................................................... ix
LIST OF FIGURES ........................................................................................................................................... x
ABSTRACT ....................................................................................................................................................... xi

CHAPTER 1: INTRODUCTION ...................................................................................................................... 1

CHAPTER 2: LITERATURE REVIEW ............................................................................................................. 8

2.1 Introduction ............................................................................................................................................... 8

2.2 The Audit Review Process ..................................................................................................................... 8

2.2.1 Characteristics Affecting Reviewer Detection and Correction of Errors and Biases ......................... 9

2.2.1.1 Preparer and Reviewer Characteristics that Influence Reviewer Performance ............................. 10

2.2.1.2 Workpaper Characteristics that Influence Reviewer Performance ............................................. 14

2.2.2 The Influence of Review on Preparer Behavior and Performance ..................................................... 18

2.2.3 Contextual Features of Review Note Closing Phase ........................................................................ 23

2.3 Dysfunctional Audit Behaviors ........................................................................................................... 23

2.4 Framing .................................................................................................................................................... 26

2.4.1 Equivalency Framing ......................................................................................................................... 27

2.4.2 Emphasis Framing .............................................................................................................................. 29

2.4.3 Framing in Accounting and Auditing Literature ............................................................................. 32

2.5 Time Delays ............................................................................................................................................ 34

2.5.1 Experimental Research on Time Delays ............................................................................................ 34

2.5.2 Attribution Theory and Negative Affect ............................................................................................ 39

2.5.3 Time Delays in the Audit Review Process ......................................................................................... 42
CHAPTER 3: HYPOTHESES DEVELOPMENT AND RESEARCH METHOD

3.1 Development of Hypotheses

3.1.1 Audit Review Process Effectiveness and the Review Note Closing Phase

3.1.2 Emphasis Framing and Review Note Rationale

3.1.2.1 Documentation vs. Conclusion Rationale

3.1.3 Reviewer Delay

3.1.4 The Interaction of Review Timeliness and Review Note Frame

3.1.5 Dysfunctional Preparer Behavior During Review Note Closing Phase

3.2 Research Method

3.2.1 Participants

3.2.2 Experimental Task

3.2.2.1 Development of Computerized Case

3.2.2.2 Task Procedure

3.2.2.3 Seeded Errors

3.2.3 Independent Variables

3.2.4 Pre-Testing of Review Note Frame

3.2.5 Dependent Variables

CHAPTER 4: RESULTS

4.1 Introduction

4.2 Manipulation Checks and Post-Experimental Questions

4.2.1 Review Timeliness

4.2.2 Review Note Frame

4.3 Testing of Statistical Assumptions
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4</td>
<td>Hypotheses Testing</td>
<td>89</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Introduction</td>
<td>89</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Number of Evidence Items Examined</td>
<td>89</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Time Spent Closing Review Notes</td>
<td>92</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Number of Errors Found</td>
<td>95</td>
</tr>
<tr>
<td>4.4.5</td>
<td>Over-Documentation of Procedures</td>
<td>98</td>
</tr>
<tr>
<td>4.4.6</td>
<td>Controlling for Demographic and Other Post-Experimental Variables</td>
<td>100</td>
</tr>
<tr>
<td>4.4.7</td>
<td>Results of Analyses with Time Outlier Omitted</td>
<td>100</td>
</tr>
<tr>
<td>4.5</td>
<td>Summary of Hypotheses Testing</td>
<td>101</td>
</tr>
</tbody>
</table>

CHAPTER 5: CONCLUSIONS, LIMITATIONS, AND IMPLICATIONS ............ 103

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>103</td>
</tr>
<tr>
<td>5.2</td>
<td>Conclusions</td>
<td>103</td>
</tr>
<tr>
<td>5.3</td>
<td>Limitations</td>
<td>106</td>
</tr>
<tr>
<td>5.4</td>
<td>Implications</td>
<td>108</td>
</tr>
</tbody>
</table>

LIST OF REFERENCES ................................................................... 112

FIGURES .................................................................................. 122

APPENDIX: EXPERIMENTAL INSTRUMENT ........................................ 125

TABLES .................................................................................... 202

VITAE ..................................................................................... 216
LIST OF TABLES

1. Participants By Cell .................................................................202
2. Sample Demographic Data ..........................................................203
3. Interpretation of Underlying Directive (Pre-Test One) ......................204
4. Perceptions of Frame Manipulation (Pre-Test One) ..........................205
5. Pretest Two: Analysis of Reviewer Expectation by Frame ..................206
6. Timeliness-Related Post-Experimental Questions by Group .................207
7. Frame-Related Post-Experimental Questions by Group ......................208
8. Number of Evidence Items Examined (Results of Hypotheses Testing) ....209
9. Time Spent Closing Review Notes (Results of Hypotheses Testing) .......210
10. Time Spent Closing Review Notes (Robustness Test One) .................211
11. Time Spent Closing Review Notes (Robustness Test Two) ..................212
12. Total Errors Found (Results of Hypothesis Testing) ........................213
13. Over-Documentation of Procedures (Measures One and Two) ...............214
14. Over-Documentation of Procedures (Measure Three) ........................215
LIST OF FIGURES

1. Expected Interaction of Review Timeliness and Review Note Frame on Preparers’
   Effort and Performance.................................................................122
2. Instrument Flow .............................................................................123
3. SPSS Q-Plot of Time ......................................................................124
ABSTRACT

Closing Review Notes: The Effect of Reviewer Delay and Review Note Frame on Audit Workpaper Preparers’ Effort and Performance

Audit review is a sequential, hierarchical process utilized by public accounting firms to detect and correct errors in the workpapers. In this process, workpaper preparers submit their work to more experienced auditors who, in turn, provide comments on the workpapers through written review notes designed to ensure the appropriateness of procedures performed and conclusions drawn. For the review process to succeed as a quality control mechanism, any issues or questions identified by the reviewer must be resolved and documented in the workpapers. Prior research has examined a number of important aspects of the review process. However, it has not addressed whether an issue or error identified through review ultimately gets corrected, and there may be certain contextual factors that enhance or interfere with this process. My study extends the audit review literature by examining this critical phase of the process, the review note closing phase. Through the use of an experiment, I examine the effect on this phase of two contextual factors within the reviewer’s control: the focus, or frame, of the rationale provided in the review note and the timeliness of the review. I find that both the timeliness of the review and the frame of the review note significantly impact the number of evidence items examined by preparers. A timely review also significantly increases the amount of time spent by the preparer. Importantly, the frame of the review note and review timeliness interact to affect the number of errors detected by the preparer in such a way that the combination of a timely review and conclusion-framed review notes lead to
superior performance. In addition, I find that reviewer delay leads to proportionally
greater dysfunctional behavior by the preparer (e.g., documentation of work not
performed). This study contributes to psychology research by examining the impact of
emphasis framing and time delays in a time sensitive, organizational setting. Further,
audit firms should be able to use the results of this research in their training to make
reviewers aware of the impact of their choices on preparer performance, the underlying
workpapers, and ultimately, the quality of the audit.
CHAPTER 1: INTRODUCTION

This study examines a previously unexplored stage in the audit review process, responding to and closing review notes, to determine how two contextual factors (instructional rationale and reviewer delay) affect the effort and performance of audit workpaper preparers. Audit review is a sequential, hierarchical process utilized by public accounting firms to detect and correct errors in the workpapers. Through the review process, more senior members of the audit team evaluate the work of their subordinates (Rich, Solomon, and Trotman 1997a). These more senior members (hereafter referred to as “reviewers”) provide comments or “review notes” that require follow-up by the workpaper preparer (hereafter, “preparer”). The review note closing stage occurs when the workpaper preparer addresses or “closes” these review notes (Roebuck and Trotman 1992). While a substantial body of research on the audit review process is developing, prior research has not investigated this phase of the audit. Thus, the extant literature cannot speak to whether, after the identification of an issue and creation of a review note by a reviewer: (a) the issue eventually gets resolved, (b) the workpapers, financial statements, and opinion reflect this and, therefore, (c) the review process succeeds as a quality control mechanism. That is, if the preparer does not adequately resolve the issue, and if the way in which he or she responds to the reviewer does not signal that the issue remains unresolved, then the review process fails as a quality control mechanism. I investigate factors within the control of the reviewer that may affect preparer effort and performance while closing review notes, and consider the implications for the overall quality of the audit.
Practitioners, regulators, and researchers recognize the importance of the review process as a quality control mechanism and as a method for training and inducing accountability in subordinates (AICPA 1979; Roebuck and Trotman 1992; Asare and McDaniels 1996; Rich, Solomon, and Trotman 1997a and 1997b; Brazel, Agoglia, and Hatfield 2004; Nelson and Tan 2005; Miller, Fedor, and Ramsay 2006). Prior research has examined whether a reviewer detects an error in the preparers’ work, the amount of effort a reviewer expends, the type of review note a reviewer leaves, and whether a reviewer anticipates stylization attempts by a preparer (Ballou 2001; Gibbins and Trotman 2002; Tan and Trotman 2003). However, research has not examined the next stage in the review process in which the preparer responds to and “closes” review notes. Understanding factors that affect preparer effort (i.e., the amount of supporting documentation examined and the time spent addressing the review notes) and performance (i.e., the results of any additional work performed) will shed light on this next stage of the audit workpaper review process and open the door for further exploration by future researchers.

Workpaper preparers perform audit procedures and document these procedures and results in the workpapers (Rich, Solomon, and Trotman 1997b). The procedures performed by junior staff members serve as the foundation of the audit opinion and, thus, errors or problems in their work could potentially lead to significant audit failures (Willet and Page 1996; Herrbach 2005). Therefore, the workpapers are passed on to reviewers (i.e., more experienced members of the engagement team) who examine and appraise the preparers’ work. Reviewers leave review notes which typically require preparers to perform additional procedures, improve documentation, or answer clarifying questions.
(Roebuck and Trotman 1992; Rich, Solomon, and Trotman 1997b). However, must accept some level of risk in relying upon the preparer’s work unless they are willing to reperform it, which is very inefficient and typically rare given the time constraints that are pervasive in public accounting (DeZoort and Lord 1997; Asare, Haynes, and Jenkins 2007). Therefore, it is important to consider how various contextual factors impact the quality of the preparer’s work throughout all stages of the audit, including review note closing.

Prior research indicates that anticipated contextual features (e.g., how the review is conducted) of the impending review can affect preparer judgment and related performance (Tan 1995; Brazel, Agoglia, and Hatfield 2004; Agoglia, Hatfield, and Brazel 2009). No research to date, however, has addressed how contextual features of the review impact preparer follow-through on reviewer comments. If the mere anticipation of review-related contextual features affects preparer performance in the initial preparation of the workpapers, it is likely that there will be contextual features of the review note closing process that will affect the performance of preparers while closing review notes. This study aims to provide some exploratory evidence regarding the effect of two contextual features of the review note closing stage that have the

---

1 In this study, I focus on the way in which evidence-oriented review notes (i.e., notes in which the reviewer identifies that additional evidence needs to be examined) are framed. Alternatively, presentation-oriented review notes relate to matters of physical workpaper presentation (e.g., “This explanation is not necessary on this [working paper], since it is noted in further detail [elsewhere]. Simply cross reference the decrease to the explanation.” (Ballou 2001, p. 37)). Thus, they leave little room for discretion on the part of the preparer and are generally considered of less importance than evidence-oriented review notes (Ballou 2001).
potential to affect preparer performance. Understanding which contextual features improve or impair preparer performance at this stage of the audit may impact audit efficiency if they lead to fewer rounds of review or the use of ineffective testing. Contextual features may also impact the overall quality of the audit. For example, if the reviewer phrases the review note in a less effective way or takes an excessive amount of time to return the workpapers to the preparer, the preparer’s performance could be negatively affected (e.g., failure to detect an error). If the reviewer does not perceive these negative effects (e.g., the reviewer is not aware of the undetected error), then audit quality may be reduced (e.g., the error remains undetected, the financial statements are not adjusted, and the implications of that error for the overall audit are not considered).

In this study, I manipulate two factors that are unique to this phase of audit review and are controllable by the reviewer: (1) the review note rationale and (2) the timeliness of the review. I define review note rationale as the supplementary advice or guidance contained in a review note which may be provided in order to explain why the additional work is necessary. A review note written to address a particular issue (i.e., an underlying directive such as “Please check to see that write-offs were properly authorized.”) may, and often does, contain different rationales (Roebuck and Trotman 1992). Specifically, the rationale is manipulated as having either a conclusion focus (i.e., making sure an appropriate conclusion is reached regarding management’s assertions) or a documentation focus (i.e., ensuring the defensibility of the documentation in the workpapers will withstand ex-post scrutiny). I manipulate the timeliness of the review as either timely or delayed; I define reviewer delay as the length of time between when the
preparer expects to receive the reviewed workpapers back from the reviewer and when he or she actually receives them.

The first variable, the review note rationale, responds to a call by Rich, Solomon, and Trotman (1997b) to examine review process gains and losses based on the type and style of review notes. Specifically, I examine whether the rationale used by a reviewer frames the same underlying directive in a way that differentially affects preparer performance. The results of this study contribute to the emphasis framing literature by examining the extent to which the effect of an emphasis frame will extend beyond a person’s likelihood assessments and preferences, and will, in fact, affect the effort and effectiveness of an individual’s task performance. With regard to reviewer delay, my study attempts to inform the psychology literature by examining the effects of time delays on a “time sensitive” hierarchical organizational culture, as time budgets and time deadlines are pervasive in public accounting firms (Otley and Pierce 1996; Houston 1999; DeZoort and Lord 1997; Sheldon, Thomas-Hunt, and Proell 2006). In addition to responding to calls from the extant literature, my study can benefit practitioners as well. Because both of these variables are, to some extent, under reviewers’ control, audit firms can use this research to help make reviewers aware of how their specific actions might impact the quality of their preparers’ work.

Using a 2 x 2 (review note rationale by review timeliness) between-subjects design and a computer-based instrument, I provide participants with a set of workpapers containing review notes. The instrument places participants in the role of having just received the workpapers back from their reviewer, with their task being to close the review notes. In order to retain as much realism as possible, the instrument contains
resources that participants could normally access during the audit, such as firm guidance, previously obtained support (e.g., credit memos that had already been obtained from the client), and the audit program. It also allows participants to “interact” with key client personnel in order to investigate transactions and obtain and view additional supporting evidence. The case contains seeded errors (e.g., a credit memo that has not been properly authorized) and the participant is asked to document in the workpapers any issues or errors they encounter.

I measure preparer effort and performance by the amount of evidence examined, the time spent addressing the review notes, and the number of exceptions noted. I expect documentation-focused rationales (hereafter, documentation-framed review notes) to cause the workpapers themselves to be more salient in the mind of the preparer than the actual conclusions they support. I anticipate review notes with conclusion-focused rationales (hereafter, conclusion-framed review notes) to enhance the accessibility and importance of the objective to ensure the appropriateness of the management assertion in question. I believe this focus on arriving at the correct conclusions, rather than demonstrating completion of a process, will result in greater preparer effort and performance than with documentation-framed review notes. I also anticipate that reviewer delay will negatively impact preparer performance because time delays have been found to induce a state of negative affect in the person experiencing the delay (which can have harmful consequences on job performance) and may cause the preparer to perceive the review notes as less important (Fox and Spector 1999; Blount and Janicik 2002; Sheldon, Thomas-Hunt, and Proell 2006). Further, I predict that the review note frame and reviewer delay will interact to affect preparer performance such that
conclusion framing will have a greater impact when the review is received in a timely manner.

As predicted, I find that a timely review leads to significantly more effort by the preparer as measured by the number of evidence items examined and the amount of time spent closing the review notes. Additionally, a conclusion-frame results in more effort by the preparer as measured by the amount of time spent. The timeliness of the review and the frame of the review note interact to affect the number of errors detected by the preparer in such a way that the conclusion-frame has a greater impact when the review is timely. Contrast tests show one fairly consistent finding across all three aforementioned dependent variables is that the effort and performance of preparers who receive conclusion-framed review notes in a timely manner tends to be significantly better than all other conditions. Lastly, reviewer delay leads to proportionally more instances of dysfunctional audit behavior (i.e., documentation of work not performed). The results of my study should be of interest to practitioners and researchers and should stimulate future research on this phase of the audit, as I find that certain characteristics of the review note closing process can improve preparer effort and performance. In other words, how and when a reviewer relays a request for additional work can have consequences for the overall quality of the audit.

The remainder of this document is organized as follows. Chapter 2 presents a review of relevant prior literature. Chapter 3 presents the hypotheses motivation and research method. Chapter 4 presents the results of manipulation checks and hypotheses testing. Chapter 5 offers conclusions and implications.
CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter reviews several areas of literature in order to develop a framework for studying the effects of review note frame and reviewer delay on the effort and performance of audit workpaper preparers. The second section of this chapter examines literature concerning the audit review process, while the third and fourth sections review the framing effects and dysfunctional audit behavior literatures, respectively. Finally, the fifth section explores the extant literature on time delays and the effects of negative affect that may be induced by time delays.

2.2 The Audit Review Process

The audit review process is a quality control process involving more senior members of an audit team evaluating the work of subordinates (Rich, Solomon, and Trotman 1997a). Auditing Standards require “critical review at every level of supervision of the work done and the judgment exercised by those assisting in the examination” (AICPA 1981, sec. 230.02; Trotman and Yetton 1985). Preparers, typically junior staff members, perform and document the basic audit procedures that go on to serve as the foundation of the audit opinion (Willet and Page 1996; Rich, Solomon, and Trotman 1997b; Herrbach 2005). Any problems with the work of these staff members could potentially lead to significant audit failure (Willet and Page 1996; Herrbach 2005). For example, the Public Oversight Board’s analysis of SEC Accounting and Auditing Enforcement Releases from July of 1997 through December of 1999 found that most of the misstatements contained in the releases involved relatively routine transactions and
accounts rather than more complex transactions and accounts, with revenue and accounts receivable being misstated in 26 out of 38 cases reviewed (POB 2000). The initial audit work performed on such routine transactions and accounts is typically performed by less experienced audit staff and seniors. Thus, if the audit review process fails to detect and initiate the correction of any errors or biases in the work of these less experienced auditors, it can affect audit quality (as evidenced by the SEC Enforcement Releases) (Asare and McDaniel 1996; Nelson and Tan 2005).

2.2.1 Characteristics Affecting Reviewer Detection and Correction of Errors and Biases

Early audit research demonstrates that the audit review process reduces judgment variance and increases accuracy in comparison to judgments obtained from individual auditors (Trotman 1985; Trotman and Yetton 1985). Much of the subsequent audit review process literature focuses on identifying contextual features that impact whether a reviewer will detect errors or biases during their review of the workpapers. In their field-based analysis of workpaper review, Gibbins and Trotman (2002) note that the success of audit review is susceptible to specific characteristics of the preparer, the workpapers, the client, and the reviewer himself. This assertion is supported by other research examining the audit reviewer’s effectiveness. The literature suggests that the audit reviewer does not always succeed at detecting or correcting preparer errors or biases.

The following two sections will discuss (1) preparer and reviewer characteristics and (2) workpaper characteristics that impact reviewer performance. Characteristics of the client that impact reviewer performance will be discussed within these two sections, as studies typically have manipulated client characteristics at the same time as either preparer/reviewer-related or workpaper-related variables.
2.2.1.1 Preparer and Reviewer Characteristics that Influence Reviewer Performance

Preparer-related factors that have been found to impact the performance of the audit reviewer include the reviewer’s level of familiarity with the preparer, preparer reputation, and perceived preparer risk (Asare and McDaniel 1996; Tan and Jamal 2001; Rich 2004; Asare, Haynes and Jenkins 2007). In addition to preparer characteristics, these studies investigate the impact of client (i.e., client risk) and workpaper characteristics (i.e., structure of the audit program). I choose to classify them in this section because these studies focus more heavily on the impact of preparer characteristics on reviewer performance. Reviewer characteristics that may impact the review, such as experience and ability (Ballou 2001), have also been studied and are discussed in this section as well.

One study that examines both preparer and workpaper characteristics is Asare and McDaniel (1996). In their study, they manipulate familiarity with the preparer by asking reviewers to assume that they either are familiar with the preparer (high familiarity) or that the preparer is from another office (low familiarity). In addition, they manipulate task complexity by varying whether the workpapers were prepared using a structured (low complexity) or unstructured (high complexity) audit program. They study the effects of these variables on the extent to which the reviewer reperforms the work of the preparer, the number of improperly classified audit items detected, and the number of conclusion errors. They find reviewers of familiar preparers are significantly less likely to reperform the work of the preparer but that overall reperformance rates were low (14%
versus 21%). The authors hypothesized, but did not find, that reviewers of familiar preparers would detect fewer classification errors than reviewers of unfamiliar preparers. Using the number of seeded conclusion errors detected as the dependent variable, the researchers find a significant interaction effect of their two independent variables. Specifically, when faced with a complex task, reviewers detect more conclusion errors when reviewing the work of familiar preparers than unfamiliar preparers; however, performance does not differ between the two types of preparers when the task is routine. Also, reviewers of familiar preparers find more conclusion errors on a complex task than on a routine task, but reviewers who are unfamiliar with the preparer find more conclusion errors on routine task than on a complex task.

Using interacting teams of audit seniors and managers, Tan and Jamal (2001) find that, under certain circumstances, managers’ evaluations of seniors will be influenced by the seniors’ reputation. Specifically, average managers will evaluate memos written by outstanding seniors more favorably than memos written by average seniors only when the managers know the identity of the senior. These findings suggest that managers may themselves be subject to evaluation biases during the review process. That is, they may anchor their evaluations on the prior performance or reputation of the preparer. While this paper does not examine whether or not a manager fails to detect or correct errors in the workpapers, the results suggest that at times some managers may rely too heavily on the reputation of the preparer. A manager who over-relies on the reputation of the

---

2 Reperformance was measured by asking participants, “What percentage of the audit staff’s work did you reperform?”

3 Both manager and senior participants are classified as either “outstanding” or “average” based on the firm’s performance evaluation system.
preparer may be less likely to detect or correct errors or biases in the preparer’s work when they are present.

More recently, Rich (2004) examines reviewers’ elaboration activities, which relate to the stage in which the reviewer assesses the strengths and weaknesses of the preparer’s work. He finds that reviewers exhibit greater elaboration activities when the perceived probability of preparer error is high than when it is low. Building on this work, Asare, Haynes, and Jenkins (2007) manipulate client risk (the risk that the client’s assertions are materially misstated) and workpaper preparer risk (the risk that material misstatements are not identified, investigated, and resolved by the workpaper preparer) to determine the effect of these two contextual variables on workpaper reviewer effort and accuracy. They use client engagement history (e.g., number of audit adjustments found on previous audits) as a proxy for client risk and preparer competence (e.g., ranking in comparison to her peers) as a proxy for preparer risk. Using the amount of time spent reviewing the workpapers as a measure of reviewer effort, they find a main effect for client risk but find no effect for the workpaper preparer risk variable and no interaction effect. Using the number of workpaper errors identified as a measure of reviewer accuracy, they find that reviewers in the condition in which both types of risk are high are significantly more accurate than reviewers in the high client risk/low preparer risk condition (the number of errors identified in the low client risk/low preparer risk condition did not differ significantly from the low client risk/high preparer risk condition). This is despite the fact that the number of errors presented in the workpapers was constant across conditions. The authors note that their results suggest that reviewers may allow low preparer risk to compensate for high client risk. This result can have
negative implications for audit review effectiveness, especially if the reviewer fails to recognize variables that may increase preparer risk (e.g., time pressure, workload pressure).

Ballou (2001) measures the impact of reviewer characteristics to determine whether they are associated with the number and type of review notes generated by the reviewer. The variables measured that were found to impact the generation of review notes include reviewer judgments about the acceptability of the workpapers, reviewer judgments about conscious information processing, general experience, specific experience, ability, and effort. Task-environment perceptions, confidence in judgments, and concerns about effectiveness versus efficiency were not found to impact the number and type of review notes. His findings indicate that reviewers’ judgments about the acceptability of the workpapers are negatively associated with the number of evidence-oriented provided. In addition, some measures of conscious information processing (the extent to which the reviewer considered alternative explanations and aspects of the review process) were associated with more evidence-oriented review notes. Surprisingly, general experience (number of months experience at their current firm) was negatively associated with the number of evidence-oriented review notes generated. The extent to which the reviewer followed a step by step process (classified as a type of conscious information processing) was associated with the generation of documentation-oriented review notes. Reviewers who reported being concerned with risks associated with evaluating the evidence provided by clients and who reported a greater consideration of quantitative support (two other types of conscious information processing), review experience, and a self-reported measure of reviewer ability were all negatively associated
with the generation of documentation-oriented review notes. The author concludes that many of these variables should be measured for future studies of reviewer performance but also admits that caution should be taken when generalizing from his study as the reviewer characteristics were based on self-reported measures.

**2.2.1.2 Workpaper Characteristics that Influence Reviewer Performance**

Characteristics of the workpapers themselves, such as the order of evidence presentation, the subset of information documented, the format and content of justifications or conclusions provided, and the way in which they are stylized can impact the effectiveness and/or accuracy of the review (Messier and Tubbs 1994; Rich, Solomon, and Trotman 1997a; Richuittie 1999; Yip-Ow and Tan 2000; Tan and Yip-Ow 2001; Agoglia, Kida, and Hanno 2003; Tan and Trotman 2003; Agoglia, Beaudoin, and Tsakumis 2009). In addition to workpaper characteristics, some of these studies investigate the impact of reviewer characteristics (i.e., reviewer rank, reviewer sensitivity to preparer stylization attempts) on reviewer performance. These studies are included in this section because the reviewer characteristic being examined is expected to moderate the effect of the workpaper characteristic that represents the primary issue investigated in the study.

Messier and Tubbs (1994) investigate the influence of the order in which evidence is presented in the workpapers and reviewer rank on reviewer effectiveness. Specifically, they examine whether reviewers will exhibit recency effects in the same manner as preparers and whether this effect will differ based on their rank. Using senior and manager auditors as participants, they first show that managers are less prone to the recency effect than seniors when making an independent, initial judgment (i.e., when
acting as a preparer). In addition, their results suggest that audit seniors acting as reviewers are equally subject to recency effects as seniors making an initial, independent audit judgment. Surprisingly, they find a significant recency effect for managers acting as reviewers. The recency effect demonstrated by this group was not significantly different than the group of seniors making an independent, initial judgment. Thus, utilizing more experienced reviewers may not be enough to overcome the effects of recency bias in the workpapers.

Rich, Solomon, and Trotman (1997a) characterize the audit review process from a persuasion perspective. Specifically, they assert that preparers send workpapers to reviewers with the intent of persuading reviewers that the work performed, conclusions reached, and related documentation are appropriate and defensible. That is, preparers use the workpapers to manage their reputation in the minds of the reviewers. Because they are intended to persuade the reviewer, workpapers may be “stylized” to create an impression which will enhance the preparers’ reputation. Tan and Yip-Ow (2001) and Tan and Trotman (2003) build upon the work of Rich, Solomon, and Trotman (1997a) by testing the impact of preparers’ stylization attempts on reviewer performance.

Tan and Yip-Ow (2001) explore how an initial conclusion presented in the workpapers and the manner in which the preparer structures or “stylizes” workpaper evidence can influence the reviewer’s judgment. The initial conclusion regarded whether or not an account of the client’s major customer was collectible. In this condition, the workpapers either concluded that the account was collectible or that it was uncollectible. The preparer’s memo structure was also manipulated as being presented in a neutral manner or a stylized manner to emphasize (de-emphasize) evidence consistent
(inconsistent) with the conclusion. The results indicate that reviewers rely on workpaper conclusions less when the workpapers are stylized than when they are neutral. The authors conclude that reviewers are sensitive to stylization attempts by preparers but warn that they only examine one type of stylization (in which certain evidence was emphasized or de-emphasized but all of the evidence was presented). Stylization involving the selective documentation (i.e., selective omittance) of evidence is less likely to be detected by the reviewer.

Tan and Trotman (2003) also investigate the impact of stylization, as well as reviewer rank and reviewer sensitivity to stylization attempts, on reviewers’ error detection rates. The authors acknowledge that there are many ways in which the preparer may stylize the workpapers; they focus their study on situations in which a preparer is known to stylize his work to please the reviewer and, based on the reviewer preference, is expected to focus on one particular type of error (conclusion errors or documentation errors). Participants are placed in the role of taking over as audit manager on a client in which the previous manager suddenly fell ill and that they must urgently review a set of already prepared workpapers. They are told that the preparer of the workpapers they are about to review is known to pay relatively more attention to identifying errors based on the preferences of the manager with whom he or she is working. They then manipulate preparer focus by telling participants that the previous manager was known to be particularly concerned about conclusion errors or documentation errors. Participants were either managers or seniors (i.e., the reviewer rank measure). Stylization sensitivity is measured as the difference in manager effort level in reaction to the two different stylization attempts (i.e., conclusion error versus documentation error preparer stylization
attempts). The results indicate that when the preparer is not expected to focus on preventing documentation errors, senior reviewers compensate for this by detecting more documentation errors. The more sensitive senior reviewers are to the stylization the more documentation errors they detect. On the other hand, managers’ detection of documentation errors has a curvilinear relation to their sensitivity to the preparer’s stylization (i.e., their detection rates first increase then fall with increased sensitivity to the stylization). When the preparer is not expected to focus on preventing conclusion errors, manager reviewers detect more conclusion errors with increasing sensitivity to the preparer’s stylization. In contrast, seniors’ detection of conclusion errors is not related to their sensitivity to the preparer’s stylization. The authors conclude that there are limits to the extent to which reviewers can counteract preparers’ stylization attempts, even when those stylization attempts are expected by the reviewer.

In addition to the research presented above, Agoglia, Kida, and Hanno (2003) also find that reviewers in their experiment are unable to detect and adjust for ineffective preparer judgments during review. Because their study focuses on both the initial judgment and the overall team judgment, it will be further discussed in the following section. A recent follow up study, Agoglia, Beaudoin, and Tsakumis (2009), does suggest that task-specific experience may help to reduce the extent to which preparer deficiencies flow through to impact reviewer judgments. The above research indicates that the audit reviewer’s effectiveness at detecting errors in the work of preparers can vary depending on preparer, reviewer, client, and workpaper characteristics. If such contextual features impact the effectiveness of the reviewer it stands to reason that various preparer, reviewer, client, and workpaper characteristics may also influence how
a preparer responds, or “closes”, review notes. However, none of this prior research empirically examines what happens after an error is detected by the reviewer.

2.2.2 The Influence of Review on Preparer Behavior and Performance

In addition to serving as a quality control mechanism, the audit review process is also recognized as a method for training and providing feedback to preparers. Roebuck and Trotman (1992) examine 3,008 actual review notes prepared by managers for twenty-eight audit engagements. They find that 4% of the review notes they examine have a sole purpose of providing direct advice to the subordinate. However, they note that the majority of the review notes that they categorize as written for the purposes of asking for further explanation, better documentation, or additional audit work or follow-up (which accounts for an additional 62% of the review notes they examined) also provide direct or indirect advice and/or training for the subordinate. More recently Fedor and Ramsay (2007) investigate the role of audit review as a form of feedback to the preparer. They asked their staff-level auditors to focus on a specific recent review and answer questions designed to determine their perception of how the review influenced their subsequent performance as a preparer. The authors find that, as a result of audit review, preparers report attempting to improve performance, manage reviewers’ impression of their work, and seek additional feedback from their reviewers.

In addition to examining the audit review process as a form of training and feedback to preparers, audit review has been explored as an accountability inducing mechanism (Johnson and Kaplan 1991; Lord 1992; Koonce, Anderson, and Marchant 1995; Hoffman and Patton 1997; Tan and Kao 1999). For example, Johnson and Kaplan (1991) manipulate accountability by telling one group of auditors that their judgments
would be reviewed by researchers and firm personnel and that they would be required to
explain their judgments in a breakout session later. The other group received no such
manipulation. Results indicate that accountable auditors achieved a higher level of
consensus and self-insight in their judgments than non-accountable auditors. There was
no difference in judgment consistency between the two groups. Similarly, Lord (1992)
manipulates accountability by having the accountability group receive instructions from a
well-respected partner from their firm. The partner instructs this group that the firm will
be evaluating their individual responses. The no accountability group is told that their
responses will remain anonymous. Results indicated that the accountability group was
less likely to issue an unqualified opinion for a hypothetical client than the no
expecting a review of their work document more justifications for their audit plan than
those not expecting a review of their work.

While this research generally shows that the review process does induce
accountability in preparers, there are studies in which accountability does not lead to
improved performance. For example, Hoffman and Patton (1997) find that auditors
exhibit the dilution effect regardless of whether they are held accountable for their work.
However, accountable auditors do provide more conservative fraud risk judgments than
auditors not held accountable. In addition, Tan and Kao (1999) investigate the
interaction of task complexity, knowledge, problem solving ability, and accountability.
They find that accountability does not enhance performance in a low-complexity task.
For a moderately complex task, accountability improves performance only when
knowledge is high. For a high-complexity task, accountability improves performance
only when both knowledge and problem-solving ability are high. Thus, while accountability is an aspect of the review process that can improve performance, it does not always succeed in eliciting the optimal performance from preparers.

While the studies discussed above have typically compared the performance of preparers anticipating a review to those not expecting their work to be reviewed (a condition not representative of the audit environment), further research has examined factors that influence accountability effects (Peecher 1996; Agoglia, Kida, and Hanno 2003; Brazel, Agoglia, and Hatfield 2004; Miller, Fedor, and Ramsay 2006; Payne and Ramsay 2008; Agoglia, Hatfield, and Brazel 2009). For example, Peecher (1996) examines “justifiee” preferences by manipulating whether the participant is told that the firm prefers a credulous, objective, or skeptical perspective for their professionals while performing analytical procedures. His findings indicate that firm preferences do influence auditors’ judgments in the direction of the preference. However, when participants were faced with low integrity clients, the preference of the firm no longer had an effect.

Agoglia, Kida, and Hanno (2003) investigate the impact of different structures for preparer justification memos. They find that the format of the justification memo can affect the effectiveness of workpaper preparer judgments. In other words, how you ask a preparer to justify a conclusion can affect the conclusion itself. Further, as discussed above, their reviewers were unable to detect and adjust for ineffective preparer judgments during review. This suggests that ineffective workpapers increase the reviewer’s burden for detecting and correcting errors, as suboptimal judgments may persist through review and, in turn, negatively affect audit quality. Building on this work, Payne and Ramsay (2008) investigate how different forms of documentation (detailed workpapers versus
summary memos) affect audit preparers’ cognitive processes and performance. They find that requesting the use of summary memos is more efficient and results in better preparer memory for the evidence. However, the use of detailed workpapers results in better error detection for an authorization error, enhances pattern recognition, and provides evidence to the reviewer which may enhance pattern recognition at the reviewer level.

Similarly, several recent studies examine whether the format of the review impacts preparer performance (Brazel, Agoglia, and Hatfield 2004; Miller, Fedor, and Ramsay 2006; Favere-Marchesi 2006; Agoglia, Hatfield, and Brazel 2009). Brazel, Agoglia, and Hatfield (2004) examine the effect of review format (face-to-face review versus review through electronic communication). They find that the format of review affects preparers’ perceived accountability which, in turn, influences their workpaper performance. Specifically, preparers expecting a face-to-face review took significantly more time to prepare their workpapers than those expecting an electronic review, but that their judgments were more effective. These results appear, in part, to be due to a greater concern for workpaper effectiveness on the part of the face-to-face group. In a follow up study (Agoglia, Hatfield, and Brazel 2009), the authors examine whether reviewers discern and compensate for any differences based on review format. Their findings suggest that the impact of review mode persists through the review stage. In other words, reviewers’ going concern judgments were of lower quality when they reviewed the work of preparers expecting an electronic review than when they reviewed the work of preparers expecting a face-to-face review. Miller, Fedor, and Ramsay (2006) perform a field survey of reviewers and preparers and find that incorporating discussion of performance with written review notes enhances preparer motivation to improve
performance. However, this effect is moderated by experience. Discussion appears to diminish performance improvement for more experienced preparers. They speculate this may be due to the introduction of too much supervisory control, which Path-Goal Leadership theory suggests is detrimental to skilled and experienced employees. Also examining review related discussion, Favere-Marchesi (2006) examines whether the timing of discussion (concurrent with review versus post-review) and the reviewer’s familiarity with the preparer influences the number of plausible hypotheses generated by an audit team during preliminary analytical review. His findings indicate that post-review discussion appears to generate significantly more hypotheses than concurrent discussion and familiar audit teams generate more hypotheses than teams with members unfamiliar with each other. The result is as the author hypothesized based on prior research which indicates that disagreement within a group is greater when group members form their own judgments before there is any group discussion (Sniezak and Henry 1990).

Thus, the audit review process is used as a form of feedback and training for subordinate auditors. In addition, research suggests that contextual features of the review, such as accountability, type and form of justification, and format of the review can influence preparer performance and this can have a real impact on audit quality as preparer bias and errors may persist through the review process. Note that this research focuses on the pre-review performance of preparers and the resulting feedback from reviewers, or on measures of preparers’ self-perceptions. However, none of this research empirically examines the next stage of the review process: how preparers follow through and close the notes (i.e., how preparers respond to comments from their reviewers).
2.2.3 Contextual Features of Review Note Closing Phase

This section has reviewed the extant audit review process literature and has demonstrated that a variety of preparer-related, reviewer-related, client-related, workpaper-related, and review-related contextual features affect the quality of the review process. As this study is the first investigation of the review note closing phase of the review process, there are a wide variety of potential contextual features that could be examined. I choose to focus on review-related factors (specifically, the timing and style of review) primarily because they are embedded specifically in this phase of the audit and are, for the most part, under the control of the reviewer. This allows for greater practical contribution, as audit practitioners can take measures to train reviewers to more carefully consider their choices with regard to the timing and style of their review. Future researchers may study the effect of preparer-related factors such as experience and ability, reviewer-related factors such as reviewer rank and reputation, client-related factors such as risk and complexity, workpaper-related factors such as the structure of the audit program and the use of decision aids, and other review-related factors such as the tone and format of review.

2.3 Dysfunctional Audit Behaviors

Anecdotal evidence and survey research suggest that some auditors at the staff and senior level may engage in behavior that is referred to as quality threatening or dysfunctional (Otley and Pierce 1996; Donelly, Quirin, and O’Bryan 2003; Shapeero, Koh, and Killough 2003; Pierce and Sweeney 2006). Some dysfunctional behaviors in which auditors may engage include: failing to report material evidence, signing off on an audit step without actually performing it, and biasing the sample selection (Pierce and
In addition, many audit researchers consider underreporting chargeable hours to be a dysfunctional behavior and find it to be prominent in practice (Lightner 1981; Lightner, Leisenring and Winters 1983; McNair 1991; Ponemon 1992; Shapeero, Koh, and Killough 2003). Underreporting, or “eating time”, differs from other dysfunctional behaviors, however, in that it does not appear to threaten the quality of the current audit and is not considered to be as deviant (by audit practitioners) as other forms of dysfunctional behavior (Shapeero, Koh, and Killough 2003).

Otley and Pierce (1996) report results from a survey of audit seniors regarding potential determinants of dysfunctional behavior. They find the leadership style of management to be significantly associated with dysfunctional behavior. Specifically, they investigate the effect of two independent leadership style dimensions: consideration (which refers to a leader that fosters mutual trust and has respect for subordinate’s feelings and ideas) and structure (which relates to how a leader defines the role of subordinates with respect to attaining a goal). As expected, they find high levels of structure to be associated positively, and a considerate management style to be associated negatively, with dysfunctional behaviors. They also report that dysfunctional behaviors are positively correlated with a strong emphasis on attaining the time budget and negatively associated with the attainability of the time budget. In addition, they find that stated or implied supervisor approval of dysfunctional behaviors is positively associated with such behaviors, and the perceived effectiveness of the review process is negatively associated with them. This suggests preparers will be more likely to engage in such behaviors if they perceive they will be rewarded for them and are less likely to engage in such behaviors if they believe they will be uncovered. Finally, they find that the level of
commitment of the individual to the organization is negatively associated with dysfunctional behaviors.

In addition to time budgets and management style, personal traits appear to impact the propensity to engage in dysfunctional behaviors. Donnely, Quirin, and O’Bryan (2003) survey auditors to gain an understanding of factors that contribute to dysfunctional behavior. They focus on traits of the individual and find that auditors who possess an external locus-of-control report lower levels of self-reported performance, exhibit higher turnover intentions (i.e., desire to leave the firm), and are more accepting of dysfunctional behaviors than those with an internal locus of control. Other research suggests that low levels of moral reasoning and a relative focus on the consequences of moral behavior (as opposed to the inherent morality of such behavior) lead to more severe underreporting of chargeable time (Ponemon 1992; Shapeero, Koh, and Killough 2003). Ponemon (1992) finds peer pressure to be a significant antecedent to underreporting. Results of a survey by Shapeero, Koh, and Killough (2003) suggest that auditors of lower rank (i.e., staff and seniors) are more likely to engage in both underreporting and premature sign-off of audit steps.

In a more recent study, Pierce and Sweeney (2006) conduct semi-structured interviews with 25 audit seniors. They find that auditors perceive a low risk of dysfunctional behaviors being detected. They also show little consideration of the ethical implications of their actions in deciding to engage in dysfunctional behavior. Also, despite the fact that most firms claim to prohibit such behavior, they find a lack of explicit communication of this prohibition (or sanctions against perpetrators) to audit seniors and staff.
I seek to extend the dysfunctional behavior literature by identifying the actual (as opposed to self-reported) use of such behavior in an experimental setting.\(^4\) In addition, I investigate whether certain features of the review process (i.e., review timeliness and review note frame) exacerbate or alleviate the use of such behavior while preparers close review notes.

2.4 Framing

The way in which information is phrased can have unintentional effects on human information processing, regardless of whether such phrasing is relevant to the decision or judgment at hand (Tversky and Kahneman 1974; Nisbett, Zukier, and Lemley 1981; Takemura 1994; Frederickson and Miller 2004). A framing effect is said to occur when a change in the phrasing of the information provided alters the way individuals perceive and act upon a task (Jamal, Johnson, and Berryman 1995; Takemura 1994). A frame has been defined as, “a central organizing idea … for making sense of relevant events and suggesting what is at issue” (Gamson and Modigliani 1989, p. 3), and the act of framing has been described as selecting an aspect of a perceived reality and making it more salient in a communicating text (Entman 1993). In other words, a frame can refer to any phrasing that increases the prominence of a certain aspect of the underlying issue. Framing effects have been investigated in generic decision making contexts in the psychology literature, as well as in more applied contexts such as political science, communications, and accounting.

\(^4\) Ponemon (1992) used an experimental design in which audit participants actually underreported the amount of time they spent on a task. However, no experimental investigations of other dysfunctional behaviors that are directly linked to the quality of the current audit appear to have been conducted.
2.4.1 Equivalency Framing

Tversky and Kahneman (1981) provide the most seminal work on framing effects, specifically the framing of risky decisions. Building on their earlier work in which they developed prospect theory (Kahneman and Tversky 1979), they posit that framed information is either encoded as positive or negative and the comparison of that information to a reference point (that is judged neutral) determines whether a given outcome is deemed a loss or a gain (Tversky and Kahneman 1981; Levin, Schneider, and Gaeth 1998). For example, Tversky and Kahneman (1981) present two groups of students a scenario in which a disease is expected to kill 600 people. The first group is given a choice between adopting the following two programs: (1) Program A in which 200 people will be saved and (2) Program B in which there is a 1/3 probability that 600 people will be saved and a 2/3 probability that no people will be saved. Overwhelmingly, respondents choose the first scenario over the second (72% versus 28%). The second group of students is given a choice between adopting two different programs: (1) Program C in which 400 people will die and (2) Program D in which there is a 1/3 probability that nobody will die and a 2/3 probability that 600 people will die. Although Programs A and C are identical (as are Programs B and D) other than the former is framed in terms of lives saved (which would be encoded as positive by the participant) while the latter was framed in terms of lives lost (which would be encoded as negative by the participant), this reframing resulted in a reversal of respondents’ preferences: 22% chose Program C, while 78% chose Program D. This reversal of preference between the two groups is evidence of an equivalency framing effect. That is, even though the two
problems are logically equivalent, the difference in the way in which the problems are presented causes individuals to make different choices.

A vast amount of literature builds on the work of Tversky and Kahneman (1981). Many of these studies show a similar framing effect in different contexts (e.g., Neale and Bazerman 1985; Maule 1989; Highhouse and Yuce 1996). For example, Schurr (1987) found that bargaining teams that focus on net profits made less risky bargaining agreements than teams that focus on net losses. In contrast, other studies have focused on factors that may lessen these framing effects in risky decision making. For example, requiring participants to provide a rationale for their choices or to think about the problem before making their choices has been shown to attenuate these framing effects, and participant expertise in the subject matter relating to the choice can also reduce these effects (Fagley and Miller 1987; Takemura 1994; Levin, Schneider, and Gaeth 1998).

A number of studies inspired by Tversky and Kahneman (1981) also examine equivalency framing in settings that do not involve uncertainty as to the likelihood of the outcome, or “risky choices” (i.e., the probabilities included in the “lives saved/lives lost” scenarios introduce a level of risk into the choice) (Levin, Schneider, and Gaeth 1998). For example, Levin and Gaeth (1988) show that labeling ground beef as “75% lean” leads to higher perceptions of the quality of that beef than labeling it “25% fat”. Meyerowitz and Chaiken (1987) demonstrate that when the negative consequences of not engaging in breast self-examinations are stressed, women are more likely to engage in such behavior than when the positive consequences of performing the exams are stressed.
2.4.2 Emphasis Framing

Druckman (2001) distinguishes between *equivalency framing* and *emphasis framing*. He notes that equivalency framing occurs when logically equivalent words or phrases cause individuals to alter their preferences or behavior (see e.g., Tversky and Kahneman 1981 and Kida 1984), while emphasis framing involves emphasizing a subset of potentially relevant considerations (Druckman 2001). An emphasis frame acts by altering the importance or weight that the message recipient places on different cues or considerations, not by altering the content of the message recipient’s beliefs in and of themselves (Nelson and Oxley 1999). That is, the frame activates and enhances the accessibility of specific ideas or thoughts related to the underlying issue but does not change the content of the ideas or thoughts (Nelson and Oxley 1999). If, for instance, a person is “of two minds” about an issue due to two particular considerations, the frame may alter the relative weight of one of the considerations but does not cause the other consideration to change or cease to exist. Similarly, Druckman (2001) explains that a person’s response to different emphasis frames is due to the fact that the frames bring to bear a substantially different set of considerations to the issue at hand.

For example, Nelson and Oxley (1999) presented two groups of student participants with two different mock newspaper articles about a land development controversy. Both stories contained the *same objective facts* about the economic gains and the environmental risks involved in the development. However, one article contained an environmental frame which emphasized the importance of considering the environmental threat, while the other contained an economic frame which emphasized the importance of the economic benefits (two substantially different considerations). The
researchers measured the students’ opinions toward the land development, their beliefs about its likely consequences (both environmental and economic), and the importance they placed on these beliefs. They found that participants in the economic framing condition were significantly more likely than participants in the environmental framing condition to indicate: (1) they would vote for the proposed development, (2) that the environmental impact of the development would be positive, and (3) that the economic impact of the property development was important to their decision. The economic framing condition participants were also significantly less likely to indicate that the environmental impact of the property development was important to their decision. Because both conditions had the same objective facts related to both the environmental and economic impact of the development project, the authors argue that emphasis framing acts by providing cues as to belief importance rather than by changing the actual content of participants’ beliefs. As another example, support for increased government spending has been shown to significantly differ depending on whether it is framed as enabling poor people to get ahead in life (in which case the issue receives more support) or as resulting in higher taxes (in which case the issue receives less support) (Sniderman and Theriault 2004).

Hallahan (1999) argues that frames reflect judgments made by message creators and the frame works by limiting and defining the message, and thus shaping the inferences that message recipients make about the message. Using a database made up of 352 strategic decisions, Nutt (1998) suggests that, when responding to message frames, the concerns and needs implied in the frame provide the motivation to act; decision makers adopt the first workable plan that emerges and match that plan to the needs
expressed by the message creator. In other words, message frame recipients may adopt a satisficing strategy in which they attempt to meet the criteria required for adequacy, rather than to look for the best plan. In this vein, a frame may lead to inferences and/or behavior by the message recipient that were not intended by the message creator (Hallahan 1999).

With a few recent exceptions (McCaffery and Baron 2004; Shankar and Tan 2006), most of the emphasis framing literature has been in the context of political science, communication, or media studies. For example, Haider-Markel and Joslyn (2001) find that alternative gun frames (individual rights versus public safety) influence opinion regarding concealed handgun laws as well as attributions of blame for the Columbine shooting. Allen, O’loughlin, Jasperson, and Sullivan (1994) examine the impact of media framing on public support for the Gulf War. They conclude that the media’s ubiquitous and repetitious messages of support constructed a reality that stifled dissent and activated a “spiral of silence” which resulted in prolonged consensus in support for the war by the public. However, Druckman’s (2001) attempt to distinguish between equivalency framing and emphasis framing also attracted the attention of researchers in other literature streams. For example, Maule and Villejoubert (2007) include Druckman’s classification in their review of framing effects. In addition, McCaffery and Baron (2004) classify their experimental framing of a “child bonus” tax break or a “childless” tax penalty as both an equivalency frame and an emphasis frame. Such a frame has both a positive/negative distinction (tax break versus tax penalty, i.e., an equivalency frame) and an emphasis on different groups that would be affected by the policy (those with or without children, i.e., an emphasis frame).
2.4.3 Framing in Accounting and Auditing Literature

There is an extensive amount of accounting and auditing literature that examines framing in an assortment of different ways. This literature looks at the impact, the detection, and the intentional use of framing effects. For instance, Kida (1984) shows that auditors’ information search and use is affected by whether they are asked if a given firm will remain viable or if it will fail. Although auditors in both conditions appeared to attend to evidence indicative of failure, auditors in the viable condition tended to focus more on evidence suggesting the company would remain viable than those in the fail condition. McMillan and White (1993) provide evidence that auditors who choose to evaluate evidence under an error frame (i.e., that ratio fluctuations may be caused by errors in the financial statements) react more strongly to both confirming and disconfirming evidence than auditors who evaluate the same evidence under an environmental frame (i.e., one which suggests the ratio fluctuations are due to environmental factors). Further, Jamal, Johnson, and Berryman (1996) examine whether Kahneman and Tversky’s (1986) proposal that frames can be detected by transforming a problem into a “standard representation” (i.e., a one error solution in this particular setting) will assist auditors in detecting a frame used by management to mask fraud. They classify auditors who propose one standard hypothesis to interpret all inconsistencies detected in the financial statements as using a standard representation. They find that auditors using a standard representation are better able to detect frames constructed by management to mask fraud than auditors using multiple representations (i.e., more than one hypothesis to interpret each individual inconsistency).
In a managerial accounting setting, Lipe (1993) demonstrates that individuals will more favorably evaluate the performance of an employee responsible for certain expenditures when those expenditures are framed as a cost than when they are framed as a loss. Amernic and Robb (2003) look at how “the quality of earnings” can be used as a framing device by accounting educators to promote the non-technical aspect of the Intermediate Accounting Curriculum (i.e., more focus on concepts such as earnings quality, less focus on standards and preparation of GAAP-based accounting information). They report their personal experiences teaching intermediate accounting and suggest that, based on student questionnaires and course evaluations, students react more favorably to their “quality of earnings frame” than to a more traditional teaching approach. More recently, Shankar and Tan (2006) show that audit workpaper preparers engage in evidence framing by using language to highlight evidence consistent with their documented conclusions and downplaying evidence inconsistent with their conclusions.

The above research reviews the impact of emphasis framing by the media, politicians, accountants, instructors, and others on the opinions of the public or individuals. This literature shows that an emphasis frame, whether intentional by the message framer or not, can impact the opinions and preferences of the public or individuals. In addition, Nutt (1998) provides some survey evidence that emphasis framing can impact strategic decision making. However, there does not appear to be much experimental evidence regarding whether the effect of an emphasis frame (in which the same directive is issued, but simply with a different rationale for why the work should be performed) will extend beyond a person’s likelihood assessments and preferences, and will, in fact, affect the effort and effectiveness of an individual’s task performance.
2.5 Time Delays

Blount and Janicik (2001) review a broad body of literature to introduce a model of how people perceive and evaluate time in organizations. They characterize time delays as violations of temporal expectations. As they explain, people naturally create expectations about when events will occur or outcomes will be realized. When these temporal expectations are not fulfilled, they try to make sense of who or what caused the negative outcome (Blount and Janicik 2001). Thus, the way in which time delays are experienced will depend partly on the way in which the delayee attributes the delay.

2.5.1 Experimental Research on Time Delays

Social science research examining the way in which people understand time, value time, and react to time delays is broad and encompasses many different literature streams (Blount and Janicik 2001). For example, intertemporal choice studies straddle both psychology and economic research. This literature stream demonstrates that, in general, people prefer a positive outcome or event sooner rather than later and discount the value of an outcome or event for which they must wait (Keren and Roelofsma 1995). One such study, Ainslie and Haendel (1983) provide participants with the option of receiving a $1,000 certified check that could be cashed in one week or a $2,000 certified check that could not be cashed until more time had passed. The participants are asked how long the delay (in receiving the $2,000 check) would have to be before they would prefer the $1,000 check. The mean response was 31 days, which suggests an extraordinarily high annualized interest rate.

Although the general findings on intertemporal choice suggest that people prefer outcomes sooner rather than later, some of this research demonstrates instances in which
this general finding is violated. For example, in a study related to prospect theory and mental accounting, Kahneman and Tversky (1984) find that 68% of participants are willing to travel 20 minutes longer to save $5 on a $15 calculator, but only 29% of participants are willing to spend the additional time and effort to save $5 on a $125 calculator. While intertemporal choices differ from interpersonal time delays, the general results of these studies show us that people do not like to wait. In fact, in some instances, they often are willing to pay a premium to avoid waiting. However, as illustrated by Kahneman and Tversky (1984), the impact of the delayed gratification on a person’s choice can be dependent on contextual features of the problem or issue at hand.

Another stream of literature examines time delays, or response latencies, in verbal communication. Long pauses in speech have been found to indicate thoughtfulness or cognitive effort, increase perceptions of deceit, and decrease perceptions of credibility (Siegman 1979; deTurk and Miller 1985; Engstrom 1994). Building on this literature, Moon (1999) examines the impact of time delays in computer-mediated communication. In this study, student participants are asked to complete a problem-solving task involving rankings. They are given the opportunity to communicate with another individual over a computer network regarding the task after they complete their initial ranking. They are told that they should not expect delays in sending or receiving information due to the computer (and took other measures) to try to ensure that they would attribute any delay that they experienced to the collaborator. Moon manipulates the time delay as a short (0 to 1 second), medium (5 to 10 seconds), or long (13 to 18 seconds) period of time.

---

5 The other measures taken to ensure that the participants would attribute the delay to the collaborator included the following: (1) the messages appeared letter by letter; (2) the typing was unevenly timed; (3) the messages contained minor typographical errors; (4) the typing included some back-spacing and retyping; and (5) the messages contained informal language.
between when the participant clicks to send his/her message to the collaborator and when the collaborator’s typing begins to appear on the screen. He measures persuasion as the difference between the participant’s final ranking and the ranking suggested by the collaborator, with lower values indicating a higher level of persuasion. He also uses post-experimental questions to compute a measure of source credibility and the perceived quality of information obtained from the collaborator.

As expected, the results show that participants who experienced the medium time delay were significantly more persuaded by the collaborator than those who experienced either the short or long delays. Moon speculates that this result is due to the short delay being perceived as reflecting a lack of thought and the long delay being perceived as annoying and a potential indication of incompetence or deception. In addition, source credibility and information quality scores are significantly lower in the long delay condition than in the medium delay condition. Neither condition was significantly different on these measures than the short delay condition.

In a more recent study, Sheldon, Thomas-Hunt, and Proell (2006) use a similar task to examine whether the status of the collaborator moderates the impact of a perceived time delay. Using undergraduate student participants, they manipulate perceived time delay by telling the participants they will receive a response from their collaborators in either 5 minutes (perceived delay condition) or 15 minutes (no perceived delay condition). All participants receive the rankings suggested by the collaborator in 15 minutes. They manipulate partner status by telling the participant that the collaborator has either a lot of task-relevant background experience (high status condition) or none at all (low status condition). In addition to a measure of influence similar to the persuasion
measure used by Moon (1999), they post-experimentally measure perceived partner competence.

They find no significant time delay effect using the competence measure for the low status group but find that high status delayers are perceived to be more competent than high status non-delayers. The authors interpret this finding in light of previous literature which suggests that violations, or deviant behavior, committed by individuals of high status are more welcomed and accepted than those committed by individuals of low status. Despite the difference in perception of competence detected by Sheldon, Thomas-Hunt and Proell (2006), the status of the delayer does not significantly impact the amount of influence the collaborator is able to exert upon the participant (i.e., there is no effect of status on influence when there is a perceived delay). Participants in the low-status delay group are, however, significantly less influenced by the input provided by their collaborator when there is a perceived time delay than when there is not. In summary, the time delay increases perceptions of competence for high status delayers but this does not translate into any differences in the amount of influence the delayer was able to exert. The time delay does not impact perceptions of competence for low status delayers, however it does decrease the amount of influence the delayer is able to exert.

Sheldon, Thomas-Hunt, and Proell (2006) speculate that the status results discussed above are due to the use of inexperienced, student participants. Therefore, they perform a second experiment using a similar scenario (and similar measures) with graduate business students as participants. In contrast to the first study, they find that participants paired with high-status delayers perceive their partners as marginally less competent than those paired with high-status non-delayers. They attribute this to the fact
that their participants are experienced and that the scenario is embedded in an organizational context. That is, they assert that the participants have more realistic experiences from which to draw to help interpret the delay than the participants in the first experiment. As with the first experiment, there is no significant difference in the amount of influence realized by a high-status collaborator based on the presence or absence of a time delay. The authors therefore conclude that having a high status significantly ameliorates the negative effects of time delays. However, they also acknowledge that their study has several limitations and specifically speculate that a time sensitive organizational culture might serve to exacerbate the extent to which time delays are perceived as expectancy violations.\textsuperscript{6} This may enhance the negative reaction to time delays (regardless of the status of the collaborator).

The above research indicates that, under certain circumstances, time delays can impact the delayee’s impression of the delayer and the influence exerted by the delayer. There is mixed evidence that the status of a collaborator may moderate the effect of a time delay, though it remains to be seen whether any observed effects will hold using a realistic task embedded in a time sensitive organizational culture in which the participants have specific experience.\textsuperscript{7} In addition, the experimental studies examined above do not

\textsuperscript{6} The limitations acknowledged by Sheldon, Thomas-Hunt, and Proell (2006) include: (1) their findings may not be generalizable beyond work groups that work together in the same physical space, (2) organizational factors not captured in their setting may dampen or amplify the results (specifically, a “laid back” organizational culture may reduce the extent to which time delays are perceived negatively and a “time sensitive” organizational culture may increase the extent to which time delays are perceived negatively), (3) familiarity with the collaborator might impact reactions to time delays, (4) the operationalization of status is limited, and (5) the task used may differ from organizational tasks in ways that lead to different results.

\textsuperscript{7} Although Sheldon, Thomas-Hunt, and Proell (2006) describe their participants for the second experiment as experienced, they did not appear to measure experience, and it is very likely that at least some of these second experiment participants went on to pursue their graduate degrees immediately after completing their undergraduate degrees. In addition, the task used in this second experiment involved software designers
examine the impact of time delays on such variables as effort and accuracy of performance. 

2.5.2 Attribution Theory and Negative Affect

As mentioned earlier, the evaluation and impact of a time delay will be affected by the way in which the delayee makes sense of, and attributes blame for, the delay (Blount and Janicik 2001). Weiner’s (1985) seminal review of attribution theory literature explains that the evaluation of a negative experience, such as a time delay, will be evaluated on several causal dimensions which will determine the emotional response to that outcome, as well as the subsequent behavior. Although many categories and sub-categories of causal dimensions have been identified in the attribution theory literature, the three most common (or “classic”) causal dimensions are locus, stability, and controllability (Weiner 1985). Locus refers to whether the outcome in question is perceived to be a result of factors within the person responsible for the outcome (internal designation) or within the environment (external designation). In attributing a cause to an interpersonal effect, such as a time delay, stability and controllability would only be of concern for an internally designated cause. In such situations, stability relates to the extent to which the cause of the outcome is perceived to be something that will fluctuate or remain constant. Controllability refers to whether the cause of the outcome is

Both Moon (1999) and Sheldon, Thomas-Hunt, and Proell (2006) investigate the impact of time delays on such variables as influence wielded by the collaborator and perceptions of competence the participant has of the collaborator. My study focuses more on the impact of the delay on the work produced by the participant (i.e., accuracy of the results of that work, amount of effort exerted to produce that work) than on the impact of the delay on the participant’s relationship to the collaborator. In addition, Sheldon, Thomas-Hunt, and Proell manipulate only the expectation of when the collaborator is expected to respond, not the length of the delay experienced. In other words, all participants experienced the same actual wait time.
perceived to be something that the perpetrator had control over or not. Take, for an example, poor performance on an exam. Such an outcome could be attributed to: poor test taking conditions or an unfair exam (external causes); a lack of aptitude (an internal, stable, uncontrollable cause); a lack of preparedness (an internal, unstable, controllable cause); or fatigue due to the flu (an internal, unstable, uncontrollable cause).

Weiner (1985) proposes “laws” to link attributional thinking and specific emotions. He notes that first (before any attribution is made) there is a general positive or negative reaction based on whether the outcome is perceived as a success or failure. Thus, based on this and the time delay literature previously discussed, a person experiencing a time delay is likely to view the delay as a violation of temporal expectations, which is a negative outcome (Blount and Janicik 2001; Sheldon, Thomas-Hunt and Proell 2006). Therefore, the delayee will experience general negative emotions in response to the delay regardless of whatever attribution he or she makes as to the cause of the delay. Specifically, the delayee is expected to experience feelings of frustration (Kelley and Michela 1980; Sheldon, Thomas-Hunt and Proell 2006).

In addition to the general negative feelings of frustration, when the delayee assigns more of the responsibility (i.e., locus) of the delay to the delayer rather than to situational factors, he or she should perceive the delay more negatively and feel greater levels of impatience (Blount and Janicik 2001). That is, the delayee will feel more negative affect when he or she attributes the delay to the delayer rather than to environmental factors. In addition, the stability of the perceived cause should affect whether or not the delayee expects future delays while collaborating with the delayer (Kelley and Michela 1980; Weiner 1985). That is, the less stable (or more of an
anomaly) the delayee believes the circumstances are which lead to the delay, the less he or she will expect subsequent delays from the delayer (and vice versa). Finally, the perceived controllability of the delay may determine whether the delayee experiences feelings of anger or pity towards the delayer (Weiner 1985). Specifically, a delayee who perceives the cause of the delay to have been under the volitional control of the delayer is expected to feel anger toward the delayer. On the other hand, if the cause is not perceived to be under the volitional control of the delayer, the delayee is expected to feel pity towards the delayer.

Organizational behavior theory predicts that a time-delay induced negative affective state can lead to frustration, the over weighting of short-term temporal goals, and can have harmful consequences on job performance (Fox and Spector 1999; Blount and Janicik 2002). Fox and Spector (1999) survey 185 corporate employees from a variety of industries in order to determine and understand causes of counterproductive work behaviors. Examples of harmful consequences that they find to be related to frustration include absenteeism, turnover, organizational aggression, and interpersonal aggression (Fox and Specter 1999). Fox and Specter (1999) also find that the form of the aggressive response taken by the employee will be strongly influenced by the perception of whether he or she can “get away with it” without getting caught.

The above research indicates that the way in which an individual attributes a negative outcome (such as a time delay) can impact the type of negative affect experienced. As Blount and Janicik (2001) argue, time delays will be affected by how the individual experiencing the delay makes sense of it (though the delayee is expected to feel negative affect regardless of the attribution made). The subsequent chapter lays out
the attribution that I expect audit workpaper preparers to make when faced with a time delay in receiving the workpapers back from their reviewer and the expected consequences of that attribution. The following section discusses extant literature on time delays in the audit review process.

2.5.3 Time Delays in the Audit Review Process

Timeliness was identified by early audit research as an important aspect of the review process (Wolf 1981; Bamber and Bylinski 1982; Wright 1985). In a survey of auditors of varying ranks, Wolf (1981) asked respondents to provide descriptions of favorable and unfavorable incidents involving an audit manager with which they had interacted. The most frequently reported unfavorable incident listed by field staff were occasions in which the audit manager failed to provide a timely review of the workpapers. Bamber and Bylinski (1982) conducted interviews with audit managers and partners and noted scheduling, time constraints, and timely reviews are all very important aspects of the audit that are fundamental for avoiding unnecessary delays in audit work. Wright (1985) surveyed audit seniors regarding performance evaluation. He found that one of the problems indicated by respondents is the timeliness of performance feedback and noted that untimely evaluations can have a negative impact on staff motivation and performance. More recently, Gibbins and Trotman (2002) conducted a survey in which audit managers were asked to list qualities of excellent and poor reviewers. Thirty-one percent of the managers surveyed referred to the timeliness of the review as an important attribute.
2.6 Conclusions

In this chapter I have reviewed extant literature from accounting, psychology, and other related fields that is relevant to the audit review process, dysfunctional audit behavior, framing, and time delays. In the following chapter, I will use this research to build hypotheses. I will also describe the method I use to test these hypotheses in the chapter that follows.
CHAPTER 3: HYPOTHESES DEVELOPMENT AND RESEARCH METHOD

3.1 Development of Hypotheses

3.1.1 Audit Review Process Effectiveness and the Review Note Closing Phase

Through the audit review process, audit firms provide feedback and training to their workpaper preparers, as well as hold them accountable for their work (AICPA 1979; Rich, Solomon, and Trotman 1997a; Agoglia, Kida, and Hanno 2003; Brazel, Agoglia, and Hatfield 2004). However, quality control is the primary purpose of the review process as public accounting firms rely on it to detect and correct biases or errors made in the judgments and decisions of less experienced auditors (Asare and McDaniel 1996; Nelson and Tan 2005). Previous research has examined the effectiveness of the review process in identifying errors in the workpapers and in preparers’ judgments (Messier and Tubbs 1994; Asare and McDaniel 1996; Asare, Haynes, and Jenkins 2007). Recent survey-based research has also examined the impact of audit review on preparers’ self-reported improvement on subsequent audits (Fedor and Ramsay 2007). However, prior research has not examined the impact of review-specific variables (such as the frame of the review note and timeliness of the review) on preparer performance during the review note closing phase. Responding to requests to advance our understanding of the sources of review process gains and losses, this study examines the practice of audit review beyond the point of error detection (Rich, Solomon, and Trotman 1997b; Tan and Trotman 2003). Specifically, I explore how review-related contextual features may impact the effectiveness of preparer performance during the correction stage of the audit (i.e., the review note closing process).
Prior research has not investigated this phase of the audit and thus cannot speak to whether an issue or error identified in the review process is corrected and, in turn, whether the review process is ultimately effective.\(^9\) As one Big 4 audit senior manager interviewed for the pre-testing phase of this study explained, “I know first hand that simply leaving the notes in no way, shape, or form is a certainty that the issue is addressed.” Preparers may misunderstand the essence of a review note or (just as is possible with initial workpaper preparation) they may act strategically in order to manage their reputations while closing the review note (Rich, Solomon, and Trotman 1997a). Depending on the preparer’s response, the reviewer may not realize the preparer has misunderstood the note or may not recognize the strategic nature of the response and, thus, the fact that the issue was not appropriately addressed may go undetected through subsequent rounds of review.

The reviewer must accept some level of preparer risk (i.e., the risk that errors are not identified, investigated, and resolved by the preparer) (Asare, Haynes, and Jenkins 2007). Reperformance of preparers’ work is considered inefficient, contrary to the purpose of review, and is rare in practice. Thus, it is advantageous to understand contextual factors that impact the way in which preparers respond to, and close, review notes. Identifying factors under the reviewers’ control that affect preparers’ performance in the review note closing process would be particularly beneficial to reviewers and could

\(^9\) For example, Ballou (2001) uses the number of review notes calling for additional evidence to be gathered as a proxy for the likelihood that the preparer will seek enough additional evidence and, thus, deems the review process effective if the reviewer identifies an area that needs additional evidence. Although it is likely that there is a positive relationship between reviewer identification of an area and additional preparer work in that area, there is no guarantee that reviewer identification of a potential issue leads to effective exploration of that potential issue by the preparer. I explore factors (specific to the note closing process) that may influence the strength of that relationship (e.g., factors that may affect the likelihood that identified potential issues are appropriately followed up on by the preparer).
assist audit firms in designing training programs and decision aids tailored toward improving the overall quality of the workpapers and the audit. One contextual factor I examine that might affect preparer performance and effort is the review note rationale. This should shed light on whether review process gains can be realized based on whether and how reviewers tailor review notes (Rich, Solomon, and Trotman 1997b).

Prior audit research and discussions with current and former auditors suggest a second factor that may affect preparer performance and effort: the timeliness of the review. As discussed in Chapter 2, timeliness was identified by early audit research as an important aspect of the review process (Wolf 1981; Bamber and Bylinski 1982; Wright 1985). Investigating review timeliness also addresses an issue raised in the psychology literature regarding the effects of time delays on “time sensitive” organizational cultures (Sheldon, Thomas-Hunt, and Proell 2006). The public accounting environment can undoubtedly be characterized as time sensitive in nature. Time budgets and time deadlines have always been pervasive in public accounting firms, but have become even more demanding in recent years due to the combination of additional internal control testing required by the Sarbanes-Oxley Act and legislation passed by the SEC which reduced the 10-K filing deadline (Otley and Pierce 1996; Houston 1999; DeZoort and Lord 1997; SEC 2005; Lambert, Brazel, and Jones 2009). Because audit managers typically work on several engagements concurrently and must adjust their priorities accordingly, some delay in returning reviewed workpapers to the preparer is inevitable (Agoglia, Brazel, Hatfield, and Jackson 2009). However, research has shown that reviewers still have discretion over how/when to conduct their reviews and can take measures (e.g., electronically reviewing a larger percentage of work, setting review time
goals for themselves, requesting additional assistance from another manager or partner) to avoid allowing the delay to become unnecessarily drawn out (Agoglia, Brazel, Hatfield, and Jackson 2009). Although there has been a good deal of research on the effect of time budgets, time deadlines, and workload pressure in the auditing literature, these studies have not examined the effect of interpersonal time delays (i.e., when one person is delayed in responding to another) in an auditing or accounting context (Otley and Pierce 1996; Houston 1999; DeZoort and Lord 1997, Agoglia, Brazel, Hatfield, and Jackson 2009).

3.1.2 Emphasis Framing and Review Note Rationale

Framing effects have been examined in political science and communication studies where a frame has been defined as “a central organizing idea . . . for making sense of relevant events and suggesting what is at issue” (Gamson and Modigliani 1989, p. 3), and the act of framing has been described as selecting an aspect of a perceived reality and making it more salient in a communicating text (Entman 1993). In other words, a frame can refer to any phrasing that increases the prominence of a certain aspect of the underlying issue. Much of the framing research in the political science and communication fields has focused on what has been termed “emphasis framing”. Emphasis framing involves highlighting a subset of potentially relevant considerations related to an issue or problem, which causes individuals to focus on those highlighted considerations (Druckman 2001).

For example, support for increased government

---

10 For a discussion of the similarities and differences between emphasis framing and what Druckman (2001) refers to as equivalency framing (which has been more frequently studied in accounting and organizational behavior settings) in which equivalent statements are worded differently, see Druckman (2001), McCaffery and Baron (2004), and Chapter 2 of this dissertation where I provide a review of the broader framing effects literature in addition to the literature on emphasis framing.
spending has been shown to significantly differ depending on whether it is framed as enabling poor people to get ahead in life (engendering more support) or as resulting in higher taxes (engendering less support) (Sniderman and Theriault 2004). With emphasis framing, preferences change because a consequentially different consideration is brought to bear on the same underlying issue (Druckman 2001). In the example above, the issue of increased government spending remains the same and, in actuality, it both enables poor people to get ahead in life and also results in higher taxes. However, emphasizing one consequence of increased government spending over another affects behavior by causing the individual to view the issue through a particular lens. Specifically, the frame selectively enhances the psychological importance, relevance, or weight accorded to a specific aspect of the issue at hand (Nelson, Oxley, and Clawson 1997; Nelson and Oxley 1999).

When creating a review note, a reviewer may include phrasing (in addition to the underlying directive or issue at hand) that communicates his or her intention in leaving the note and provides direct and indirect advice/guidance to subordinates (Roebuck and Trotman 1992). If this wording choice emphasizes one aspect of the underlying issue over some other aspect, it may act as an emphasis frame and cause the preparer to evaluate the underlying directive through the lens of that frame (Druckman 2001). Reviewers may (consciously or subconsciously) use such phrasing because the audit review process is an important part of the training, coaching, and development of auditors (Roebuck and Trotman 1992; Wiongrad, Gerson, and Berline 2000; Brazel, Agoglia, and Hatfield 2004; Miller, Fedor, and Ramsay 2006). That is, review notes serve as a form of feedback and guidance to preparers in addition to their role as a means of requesting
follow up work or documentation (AICPA 1979; Roebuck and Trotman 1992; Rich, Solomon, and Trotman 1997a).

In their examination of actual review notes prepared by auditors, Roebuck and Trotman (1992) find a substantial amount of both direct and indirect advice provided by the reviewers. While only 4% of the review notes they examined had the sole purpose of providing advice to the preparers, they find that most of the review notes that fall into the other categories also provided indirect advice and training to subordinates (Roebuck and Trotman 1992). In their study, they find that one of the most common purposes of review notes is to request additional audit work or follow-up. The additional advice or guidance (i.e., instructional rationale) contained in a review note may be provided as a way of explaining why the additional follow-up work (i.e., underlying directive) being requested is necessary. For instance, Ballou (2001 p. 37) provides the following example of a review note rationale that has an underlying directive to investigate whether gross interest expense decreased: “Not only would this corroborate the shift in income, but it will give us an idea as to whether they are doing an effective job managing assets/liabilities and interest rate risk.” The rationale explains to the preparer why they should investigate gross interest expense: because it will corroborate the shift in income and provide information on the client’s interest rate risk.

It is important for preparers to understand the nature and purpose of the procedures they perform as this may prove to be valuable to the long-term development of the auditor as both a preparer (e.g., to better anticipate and avoid similar review notes in the future) and a future reviewer (e.g., to have a stronger foundation from which to draw when he or she is reviewing workpapers) (Rich, Solomon, and Trotman 1997a;
Fedor and Ramsay 2007). However, the phrasing of this additional rationale may also act as a frame by selectively enhancing the importance, relevance, or weight of one potential purpose of the underlying directive at the expense of another (Nelson, Oxley, and Clawson 1997; Nelson and Oxley 1999). In other words, the rationale reviewers provide in order to help train the preparer may have framing effects on the effort and performance of the preparer. In highlighting one objective of the audit, a reviewer may thus reduce the relative importance of another objective in the mind of the preparer, which may lead to unanticipated effects on the quality of the current audit at hand.

3.1.2.1 Documentation vs. Conclusion Rationale

There are a number of ways a reviewer may frame a review note, as audit firms do not have pre-populated lists of review notes from which to choose and the review note is written by the reviewer in his or her own words (Ramsay 1994). Since managers show a good deal of variation in the number, type, and style of review notes they leave (Roebuck and Trotman 1992; Gibbins and Trotman 2002), it is likely that a review note with the same underlying directive may be presented with a different rationale.11

Rich, Solomon, and Trotman (1997a) describe two major classes of review process objectives: (1) reaching an appropriate conclusion regarding conformance with generally accepted accounting principles (GAAP) (i.e., a “conclusion objective”),12 and (2) ensuring that the documentation in the workpapers is defensible and will withstand ex post scrutiny (i.e., a “documentation objective”). They base their classification on

---

11 I discussed this likelihood with five current and former auditors. They all recalled receiving review notes (calling for additional evidence) that were worded in a way that emphasized both of the primary objectives of the reviewer (i.e., appropriateness of conclusions and adequacy of documentation). Prior research supports the notion that review notes can invoke different rationales for the work requested, including both conclusion and documentation oriented rationales (Roebuck and Trotman 1992).

12 Rick, Solomon, and Trotman (1997a) refer to this category as “opinion appropriateness”.
Roebuck and Trotman’s (1992) examination of actual review notes and observe that both of these review process objectives are consistent with professional standards. That is, the conclusion objective stems from the fact that forming an appropriate opinion is the objective of a financial statement audit, while the documentation objective arises because the working papers provide the record of procedures applied, tests performed, and information obtained during the engagement (AICPA 1995). In addition, there are distinct consequences associated with not meeting either of these objectives. If the audit conclusions are incorrect, then both audit risk (i.e., the risk that auditor provides an inappropriate opinion) and engagement risk (i.e., injury related to litigation and adverse publicity) are high (Messier, Glover, and Prawitt 2008). If the audit documentation is inadequate, the auditor risks receiving a negative PCAOB inspection which can lead to such disciplinary actions as public hearings, sanctions, and license suspensions and/or revocations (SOX 2002).

While both of these objectives are important to the reviewer, they are not necessarily both explicitly conveyed through the review note rationale. In the review note example from Ballou (2001) discussed above, the rationale has a conclusion focus. That is, the reviewer conveys that the underlying directive is intended to ensure that the conclusions being made about the shift in income are corroborated and, thus, appropriate. The reviewer may, however, leave a review note that stipulates performance of a procedure, but include a rationale that conveys to the preparer the importance of making sure that there is sufficient, defensible documentation in the workpapers. An example of a review note with a documentation frame might read as follows: “Please check to see
that write-offs were properly authorized. We need to make sure we document performance of all of the steps in the audit program."\(^\text{13}\)

The defensibility of the workpapers has become an even more important objective recently with the passage of Auditing Standard No. 3 (AS3), *Audit Documentation*, by the Public Company Accounting Oversight Board (PCAOB). The PCAOB issued AS3 because “the board has made documentation a priority” and noted that “clear and comprehensive audit documentation is essential to enhance the quality of the audit and, at the same time, to allow the Board to fulfill its mandate to inspect registered public accounting firms with applicable standards and laws” (PCAOB 2004, p. 320). Rich, Solomon, and Trotman (1997a) note that reviewers sometimes direct preparers to perform additional procedures not because they believe those procedures will have additional diagnostic value, but because they believe they are necessary to make the workpapers more defensible. The passage of AS3 may increase the salience of the documentation objective in reviewers’ minds and may therefore lead reviewers to more frequently frame their review notes in order to highlight this objective.

It is important to understand both the way in which a frame is perceived/understood by message recipients and how that perception determines their subsequent behavior (Maule and Villejoubert 2007). The frame of a review note may impact the preparer’s perception of the underlying directive by selectively enhancing the accessibility of, and relative weight that the preparer places on, whichever reviewer

\(^\text{13}\) This same underlying directive with a conclusion frame could read as follows, “Please check to see that write-offs were properly authorized. We need to make sure our conclusions about the authorization assertion are accurate.” Notice that, in both review notes, the preparer is being told to complete the same additional procedure; it is simply the reason behind the completion of the procedure that is different. It is also possible that a review note could have multiple frames.
objective is the focus of the frame (Nelson, Oxley, and Clawson 1997; Nelson and Oxley 1999; Druckman 2001). In other words, when the review note has a conclusion focus (documentation focus), this both primes the “appropriate conclusion” (“defensibility of the documentation”) objective and increases the relative importance of that objective in the mind of the preparer (Kinder and Sanders 1996; Nelson and Oxley 1999). The question that remains is how that internal representation might affect preparer performance in closing review notes given that the underlying directive is identical (i.e., the request for additional work is worded in exactly the same way).

A frame reflects judgments made by the creator of the message, whether or not the effect of the frame was intentional or unintentional (Hallahan 1999). The concerns and needs implied in the communication provide the message recipient with the motivation to act (Nutt 1998). Preparers responding to a conclusion-framed review note may sense a greater reviewer concern for the accuracy/appropriateness of the management assertion in question. Although the underlying directive is exactly the same, this may increase the salience of both the necessity of the additional procedure and the possibility that an actual error exists in the area being tested (Druckman 2001). When auditors believe there is a realistic possibility that an error or misstatement exists, this should affect the nature, timing, and extent of tests they perform (Peecher 1996). Preparers primed with the “appropriate conclusion” objective are likely to be more sensitive to ensuring the correctness or accuracy of the account balance or assertion being tested by the underlying directive in the review note. This may affect preparer performance by causing the preparer to take the time to gather further evidence and evaluate it more carefully. That is, a conclusion-framed review note (relative to a
documentation-framed review note) may lead to a greater level of effort and performance on the part of the preparer in closing the review note.

A preparer viewing the underlying directive through a documentation-focused lens may perceive that the reviewer is primarily concerned with the presentation of the workpapers. AS3 defines audit documentation as “the written record of the basis for the auditor’s conclusions that provides the support for the auditor’s representations” (PCAOB 2004, p. 309). A documentation-framed review note may cause the written record (i.e., the workpapers) to be more salient in the preparer’s mind than the actual conclusions which that written record support (Druckman 2001). This focus on the formalities of the audit process, while important, may cause the preparer to view the purpose of the additional audit work as demonstrating completion of the process rather than arriving at the correct conclusions (Knechel 2007). While the reviewer may not intend to leave the preparer with this impression, frames may lead to inferences that are not the intent of the message originator (Hallahan 1999).

Including a rationale that focuses on AS3-related documentation requirements may not adequately remind preparers of the underlying reason such documentation is necessary. Particularly in an organizational setting, recipients of a message frame frequently match their plan of action to the concerns or needs of the message originator, often rationalizing a plan of action that is simply “good enough” to meet or exceed particular criteria but that does not necessarily lead to the best outcome (Nutt 1998). In addition, preparers may be less “primed” to catch an issue or problem while they perform the procedures required by the underlying objective. Not really expecting to find errors while performing the procedures, the preparer may perform the procedures requested by
the reviewer more hurriedly (and less conscientiously) than a preparer receiving conclusion-framed review notes. In turn, the preparer may fail to obtain the sufficient amount of evidence, or fail to sufficiently examine the evidence obtained, to identify issues or errors he or she otherwise *should* have discovered while gathering the documentation necessary to satisfy AS3. Based on the above discussion, my first set of hypotheses are:

**H1a:** Preparers will examine more supporting evidence when closing conclusion-framed review notes than when closing documentation-framed review notes.

**H1b:** Preparers will spend more time closing conclusion-framed review notes than closing documentation-framed review notes.

**H1c:** Preparers will identify more errors when closing conclusion-framed review notes than when closing documentation-framed review notes.

### 3.1.3 Reviewer Delay

People naturally create expectations about when events will occur or outcomes will be realized (Blount and Jancik 2001). When these temporal expectations are not fulfilled, they try to make sense of who or what caused the outcome to *not* be realized in the expected time frame (Blount and Jancik 2001). Attribution theory predicts that a negative outcome, such as a time delay, will be evaluated on several causal dimensions (e.g., locus, stability, and controllability) which will determine the emotional response to that outcome, as well as the subsequent behavior (Weiner 1985).\(^{14}\) As I will explain in more detail later in this section, assuming an auditor attributes reviewer delay at least partly to the reviewer, the most relevant dimension for the purposes of my setting is

---

\(^{14}\) See chapter two for a more in-depth discussion of these causal dimensions.
controllability. The perceived controllability of the delay may determine whether the delayee experiences feelings of anger towards the delayer (Weiner 1985). Specifically, a delayee who perceives the cause of the delay to have been under the volitional control of the delayer is expected to feel anger toward the delayer.

People differ in their ability to tolerate delays, and the perceived causality of the delay will differ from person to person and within an individual over occasions (Weiner 1985; Blount and Janicik 2001). However, no matter what attribution the person experiencing a delay makes (along any causal dimension), he or she may still experience frustration if the delay is perceived as a negative outcome (Kelley and Michela 1980; Weiner 1985). For example, in an audit environment, even if the preparer attributes the delay to external forces completely beyond the control of the reviewer, to the extent that the preparer perceives the delay to be a negative outcome in response to the original work performed, the preparer is likely to experience negative affect due to the delay. Organizational behavior theory predicts that a time-delay induced negative affective state can lead to the overweighing of short-term temporal goals and have harmful consequences on job performance (Fox and Spector 1999; Blount and Janicik 2002).

Timeliness was identified by early audit researchers as an important aspect of the review process (Wolf 1981; Bamber and Bylinski 1982). Delays in formal and informal feedback to audit staff can affect their motivation and performance and result in inefficiencies (Wolf 1981; Wright 1985). According to Wolf (1981), the most frequent unfavorable incident reported by field staff was instances in which the audit manager failed to provide a timely review of the workpapers. Based on discussions with current and recent Big 4 senior auditors, review timeliness still appears to be considered by field
staff to be an important factor, as it was identified as a determinant of the amount of time and effort they devoted to closing review notes. As one senior auditor explained, “When you’ve got the stuff in your head, you’re thinking about it. If they take too long to review, you lose your train of thought and care less.” There are many particulars of the audit environment that can make reviewer delay particularly frustrating to preparers. For example: the preparer has likely switched to working on another area and now has to change his or her plan of action and train of thought to deal with the reviewed area; the preparer may have rolled onto another client and now has to juggle the demands of working for two managers and trying to obtain information from the client while working at a different location; the client may be more contentious or more difficult to deal with because they believed testing on the area in question was already completed.

I am interested in the extent to which this delay impacts the work of the preparer. There is evidence that managers may not realize the extent to which reviewer delay affects the attitude and performance of the preparer despite the fact that they were once preparers themselves. Gibbins and Trotman (2002) conducted a survey in which audit managers were asked to list qualities of excellent and poor reviewers. In contrast to the audit seniors with which I spoke in preparation for this research (who each referred to reviewer delay as an important factor in their subsequent performance), only 31 percent of the managers surveyed by Gibbins and Trotman (2002) referred to the timeliness of the review as an important attribute. This disconnect (i.e., audit seniors appear to value review timeliness more highly than audit managers) may be due to the fact that audit

---

15 During the initial process of identifying key contextual factors of this stage of the review process, I spoke with five former and practicing auditors. I asked them open-ended questions regarding what factors they felt have impacted the amount of time and effort they spent closing review notes as a preparer. All five of them identified reviewer delay as one of the more important factors.
managers typically work on several engagements concurrently, adjusting priorities
(Aagoglia, Brazel, Hatfield, and Jackson 2009), and therefore may consider occasional
time delays in returning reviewer workpapers inevitable.

On the other hand, audit staff members typically serve one client at a time and,
thus, likely view their current client as the top priority (Dodin and Eliman 1997). If the
audit manager takes a long time to review their workpapers, preparers may perceive that
the manager is busy working on other clients and adjusting priorities. The fact that it has
taken so long for the manager to review this specific work may then lead the preparer to
the assumption that the manager considers the work in question to be of a lower priority.
In addition, there is evidence that employees that are at higher levels in an organization
tend to have longer time horizons than those at lower levels (Blount and Janicik 2001).

If reviewer delay does impact preparer follow through of review notes, then there
are important implications for audit firms and managers. Although the current
environment may be associated with higher levels of workload pressure for audit
managers, reviewers ultimately have discretion over how/when to conduct their reviews
(Aagoglia, Brazel, Hatfield, and Jackson 2009). Reviewers can take measures, as
discussed previously, to avoid allowing the delay to become unnecessarily drawn out
(Aagoglia, Brazel, Hatfield, and Jackson 2009). If a delay is entirely unavoidable, the
reviewer can try to set a reasonable expectation for the preparer as to when the review
will be completed and can explain the reason for the delay.\footnote{Existing psychology theory suggests that setting an accurate expectation and providing an explanation for an extensive time delay might alleviate some of the negative effects of that delay. The focus of my study is to provide initial evidence on the impact of time delay in an audit setting. I leave it to future research to explore the effect of expectations/explanations in such settings.}
As discussed earlier, when a preparer faces an extensive time delay (i.e., significantly longer than anticipated), regardless of any attributions he or she may make, it is likely to put him or her in a negative affective state which can lead to suboptimal cognitive processing and inferior performance (Fox and Spector 1999; Forgas and George 2001; Blount and Janicik 2002). In searching for an understanding of the cause for the delay, attribution theory predicts that the preparer will consider how much control the reviewer had over the delay (Weiner 1985). In addition, Sheldon, Thomas-Hunt, and Proell (2006) find time delays cause participants in laboratory experiments to view information provided by the collaborator responsible for the delay as less important (Sheldon, Thomas-Hunt, and Proell 2006). They find that the negative effects of time delays are moderated by the status of a collaborator. That is, delays caused by high-status individuals have less of an impact on the response of delayees than a delay caused by low-status individuals. However, they admit to a limited and narrow operationalization of status in their first experiment due to the use of undergraduate student subjects (without organizational experience) placed in a simulation lacking a true organizational context (Sheldon, Thomas-Hunt, and Proell 2006). They perform a second experiment using graduate students but the amount of organizational experience that these participants have had does not appear to have been measured, and it is not evident that they actually had experience related to the experimental task. Thus, the extent to which time delays affect the effort and accuracy of performance of skilled participants in a hierarchical, organizational environment is an empirical question that has not yet been investigated. In fact, Sheldon, Thomas-Hunt and Proell (2006) recognize the need for, and call for, additional research on the way in which time delays influence behavior in
time-sensitive, organizational settings. This study extends current organizational theory by examining delays in such a setting.

As discussed above, the attribution associated with a time delay may impact the way in which the delayee reacts to the delay. Based on my discussions with current and former auditors (see footnote 14), the most frequent cause they would attribute to an extensive delay is that the reviewer is busy reviewing other areas and working on other clients and must juggle their priorities. Such an attribution has aspects that are both environmental and internal (reviewer-related). While the preparer’s personal characteristics and his or her own organizational experience will likely affect the attributional process (Blount and Janicik 2001), based on my personal experience and discussions with current and former auditors (see footnote 14), I expect most preparers to view the delay as controllable because they perceive that the reviewer is setting his or her own priorities. That is, the preparer realizes that the reason for the delay is because the reviewer is busy adjusting priorities but also makes the assumption that the work in question must therefore be of a low priority to the reviewer or he or she would have juggled it to the top of his or her priority list. This line of reasoning, while somewhat speculative due to this being the first foray into the study of reviewer delay in the auditing literature, is in line with the Sheldon, Thomas-Hunt, and Proell’s (2006) finding that information provided by a collaborator is deemed less important when there is a time delay. Thus, in addition to being associated with feelings of frustration, reviewer delay may cause the preparer to view the additional work requested in the review notes as less important than if there had not been a delay.
In summation, I expect reviewer delay to lead to frustration and possibly anger on the part of the preparer and to potentially cause the preparer to view the review notes as less of a priority than when there is no delay. This may lead to lower levels of effort and, in turn, performance on the preparer’s part. Based on the above discussion, I hypothesize the following:

**H2a:** Preparers will examine less supporting evidence when there is a reviewer delay than when there is no reviewer delay.

**H2b:** Preparers will spend less time closing review notes when there is a reviewer delay than when there is no reviewer delay.

**H2c:** Preparers will identify fewer errors when there is a reviewer delay than when there is no reviewer delay.

### 3.1.4 The Interaction of Review Timeliness and Review Note Frame

I expect the combination of a timely review and conclusion-framed review notes to result in optimal performance by the preparer. One of the variables identified by Druckman (2001) that has been found to limit the effect of emphasis framing is trustworthiness. For example, he reports results of a study that shows a larger framing effect for participants reading an article from the *New York Times* than from the *National Enquirer*. Trust involves having confidence that some other person or entity will not act in a way that results in distressing circumstances (Boon 1995). Reviewer delay is expected to leave the preparer in a negative affective state in which the preparer is frustrated and/or angry. Negative affect (e.g., anger) has been found to be associated with significantly lower levels of trust in another person even when the negative affect is not specifically associated with that other person (Dunn and Schweitzer 2005). To the
extent that the preparer is distressed/frustrated by the delay, the preparer may feel that the
reviewer has violated his or her trust and may feel negatively toward the reviewer. Thus,
the impact of the frame may be weaker when there is a reviewer delay. That is, the
rationale expressed in the review note by a reviewer providing a timely review may be
considered more closely by a preparer than a rationale supplied by a reviewer providing a
delayed review. Therefore, I expect review timeliness and review note frame to interact
in such a way that the effect of the frame will be greater when the review is timely, and I
propose the following hypotheses (see Figure 1 for a graphical representation of the
expected interaction):

**H3a:** Review timeliness and review note frame will interact such that the positive effect
of a conclusion (relative to documentation) frame on the amount of supporting
evidence preparers examine is greater when the review is timely that when the
review is delayed.

**H3b:** Review timeliness and review note frame will interact such that the positive effect
of a conclusion (relative to documentation) frame on the amount of time preparers
spend closing review notes is greater when the review is timely than when the
review is delayed.

**H3c:** Review timeliness and review note frame will interact such that the positive effect
of a conclusion (relative to documentation) frame on the number of errors
preparers identify is greater when the review is timely than when the review is
delayed.
3.1.5 Dysfunctional Preparer Behavior During Review Note Closing Phase

The audit literature provides evidence that, for the most part, auditors try to execute their roles in a conscientious manner. However, there is some anecdotal evidence and survey-based research which suggests that some auditors at the staff and senior level may engage in behavior that is referred to as quality threatening or dysfunctional (Pierce and Sweeney 2006; Donnelly, Quirin, and O’Bryan 2003; Rich, Solomon, and Trotman 1997a; Otley and Pierce 1996). Some dysfunctional behaviors in which auditors may engage include: failing to report material evidence, signing off on an audit step without actually performing it (sometimes referred to as over-documentation of audit procedures or “ghost ticking”), and biasing the sample selection (Pierce and Sweeney 2006; Rich, Solomon, and Trotman 1997a).

Research on dysfunctional behavior has primarily been conducted through the use of interviews and surveys and finds associations between self-reports of such behavior and a number of characteristics and contextual features of the audit (e.g., an external locus of control, perceived attainability of the time budget) (Donelly, Quirin, and O’Bryan 2003; Otley and Pierce 1996). However, this stream of literature has not yet experimentally demonstrated the occurrence of such behavior. The nature of my instrument allows me to detect use of at least one of these dysfunctional behaviors (i.e., over-documentation of audit procedures). In addition, no experimental research has investigated whether there are factors that can mitigate or exacerbate the use of such activities, which pose a threat to the quality of the audit. Therefore, I have two goals related to the study of dysfunctional behavior with regard to my study: (1) to examine whether experienced auditor participants will engage in the use of dysfunctional
behaviors in a realistic audit setting and (2) to examine whether there is a causal link between contextual features of the audit (specifically, features related to the audit reviewer) and the use of dysfunctional behaviors by preparers while closing review notes. Thus, I will examine whether the frame of the review note and/or the timeliness of the review can result in a lesser tendency towards dysfunctional behavior.

As outlined in chapter 3.1.3, reviewer delay may negatively impact preparer performance due to the inducement of negative affect and because the delay may cause the preparer to conclude that the area of testing is of low priority to the reviewer (Blount and Janicik 2002; Sheldon, Thomas-Hunt, and Proell 2006). Negative affect can facilitate a more externally motivated thinking style in which people focus on short-term temporal goals and seek to produce a response with the least amount of effort, using whatever shortcuts or simplifications are readily available (Forgas and George 2001; Blount and Janicik 2002). Thus, in addition to possibly concluding that the area is of low priority to the reviewer, preparers faced with an extensive reviewer delay may focus on the short-term goal of “getting the review notes closed” (as opposed to a goal of ensuring the correctness of the account balance) and be externally motivated to appease the reviewer (rather than internally motivated to do as good a job as possible).

Also, Fox and Specter (1999) find that frustration is related to organizational and interpersonal aggression and that the form of the aggressive response (e.g., behavior that may be harmful to the organization) taken by the frustrated employee will be strongly influenced by the perception of whether he or she “can get away with it”. Pierce and Sweeney (2006) note that auditors perceive there to be very low risk of a reviewer detecting a preparer’s use of quality threatening behaviors. Thus, preparers may respond
to the frustration of a time delay by engaging in dysfunctional behaviors because this is seen as a form of organizational frustration that they can “get away with”. Therefore, I propose that the timeliness of the review will impact the extent to which preparers will engage in quality threatening behaviors (specifically, for this case, in the form of over-documentation of procedures) and hypothesize the following:

**H4a:** Preparers receiving a timely review will engage in less dysfunctional behavior while closing review notes than preparers receiving a delayed review.

When preparers receive documentation-framed review notes, they may perceive the frame as a shortcut that allows them to rationalize responding to the review note using the least effort required in order to placate the reviewer. As Rich, Solomon, and Trotman (1997a) note, preparers may rationalize engaging in quality threatening behaviors “by believing their knowledge allows them to anticipate the implications of any evidence which would have been collected” (p. 493). A documentation-framed review note may cause the preparer to foresee that any additional evidence they gather is purely for documentation purposes and will not have any real consequences for the audit opinion. Therefore, I expect less dysfunctional behavior when preparers close conclusion-framed review notes than when they close documentation-framed review notes and hypothesize the following:

**H4b:** Preparers receiving conclusion-framed review notes will engage in less dysfunctional behavior while closing review notes than preparers receiving documentation-framed review notes.

One auditor interviewed for this study remarked that, when responding to review notes, an auditor might decide to just “beef up my text in my tickmark rather than go
back and make sure the testing was thorough and complete.” When asked what might influence whether he would just “beef up the text in the tickmark” as opposed to actually going back and making sure the testing was thorough and complete, he indicated that he would be influenced by the extent to which he likes the reviewer and, in some circumstances, whether he perceives the additional work will have “material consequences”.

This anecdotal evidence helps to illustrate the way in which both affect and the preparer’s perception of whether the additional work requested has material consequences may interact to determine whether an auditor engages in quality threatening behaviors (e.g., changes the tickmark to indicate additional work was done without actually doing the additional work). Consistent with this example, conclusion-framed review notes should cause the preparer to sense greater reviewer concern for the accuracy/appropriateness of the management assertion in question, and should lead to a greater sense that there may be material consequences associated with the work requested in the review note. However, as reviewer delay could lead to feelings of frustration and anger (i.e., negative affective reactions) on the part of the preparer, any positive effect of a conclusion-framed review note may be limited by the level of negative affect the preparer feels (Druckman 2001; Dunn and Schweitzer 2005). Thus, I expect the review note frame to have a greater influence on preparer dysfunctional behavior when the review is timely than when it is delayed.

**H4c:** Review timeliness and review note frame will interact such that the positive effect of conclusion (relative to documentation) frame on the propensity to engage in
dysfunctional behavior is greater when the review is timely than when the review is delayed.

3.2  Research Method

3.2.1  Participants

Participants are 69 staff and senior level auditors with an average of approximately 17 months of experience (see Table 1 for number of participants by cell and Table 2 for demographic data for the sample). The instrument involves closing review notes during the testing of accounts receivable for a hypothetical client, Sprandel Inc. Discussions with practicing audit managers revealed that audit staff of this general experience level should be familiar with the types of testing and procedures requested by the review notes in my study. I designed the review notes in such a way that some general audit testing experience (e.g., reviewing supporting evidence) is necessary, but a large amount of experience testing accounts receivable is not. Pre-testing of the review notes supported this assertion. In addition, the review notes include specific instructions to preparers regarding the additional work that must be performed and, thus, the case involves a highly structured task. Previous auditing research suggests that experience level matters less for studying performance on highly structured audit tasks (Abdolmohammadi and Wright 1987). Also, the participants do not have to make the “correct” final judgment regarding the evidence they examine, they simply need to bring any unusual or incorrect items to the reviewer’s attention.

I measure participants’ level of general audit experience (i.e., months of audit experience), experience closing review notes, experience working on manufacturing clients, and likelihood that they will be assigned to a client similar to Sprandel, Inc. As
reported in Table 2, the respective means (standard deviations) for these measures are 16.68 (13.63), 7.54 (3.09), 4.91 (3.11), and 8.23 (2.47). The audit experience for the sample ranged from 3 months to 61 months. Finally, note that (as discussed in section 4.4.6) all analyses used to perform hypotheses testing were also performed using the demographic data presented in Table 2 as covariates, and such inclusion did not significantly affect the results presented or change any of the inferences drawn. The lack of an experience effect suggests that the task was suitable and sufficiently well-structured for the less experienced participants.

3.2.2 Experimental Task

3.2.2.1 Development of Computerized Case

Participants are provided with a computerized case that I developed using Macromedia Authorware software (see appendix for description of the instrument and relevant excerpts). I developed the case based on my personal experiences as a former auditor and based on an audit program and workpapers used by a Big 4 audit firm. I used Microsoft Excel templates to create supporting invoices, shipping documents, and other supporting evidence. Throughout development of the instrument, I periodically obtained the input and advice of nine current and former practicing auditors from three of the Big 4 firms. I also pre-tested an early version of the case (without manipulations) on auditing students to ensure the understandability of the task and functionality of the software. In developing the case, I aimed to: (1) ensure that the amount of work required by the participant would be enough to generate variability on the dependent variables, (2) ensure that the type of work required by the participant would only call for general audit
experience (as opposed to experience with a specific area or client),\textsuperscript{17} and (3) use the capabilities of the software such that the task would emulate the complexities of an actual audit and would allow preparers to react as they might on a real audit.\textsuperscript{18}

### 3.2.2.2 Task Procedure

The case begins with instructions informing the participants that their task involves closing review notes on a set of workpapers that they are to assume they prepared (see Figure 2 for a concise presentation of the flow of the instrument). Participants are asked to address the review notes as they would if on an actual client and to document any additional errors (if any) they may find while closing the review notes. After reviewing the instructions, participants proceed to a background screen which provides them with general information about the client and the names and positions of the three client personnel they will be able to “communicate” with in order to obtain additional evidence. They then submit the workpapers to the reviewer and wait for the reviewer to return the workpapers.

\textsuperscript{17} This goal was necessary because the fewer restrictions I put on the type of participant I could use in my study, the more generalizable the results. Also, audit participants are a scarce resource. By designing the study to be appropriate for auditors with general audit experience, I could appeal to a larger base of potential participants.

\textsuperscript{18} It appears as though these three objectives were met. As will be shown in chapter 4, there is wide variability of responses on the dependent variables and controlling for the different measures of experience does not significantly impact the results or the inferences drawn from the results. With regard to realism, at the end of the study, participants in my experiment were asked to respond to the following question on 11-point likert scales (1 = not at all realistic, 11 = very realistic), “Based on your experience, how realistic were the 7 review notes presented in this case?” The mean response was 8.33 (SD = 2.42). Also, participants were given an opportunity to provide feedback through the use of the following open-ended question: “If you would like to provide the researchers any feedback regarding this project (any thoughts regarding the simulation, any issues navigating through the simulation, or any other thoughts regarding this project), please use the space provided below.” Most of the participants who chose to respond to this question made a point to comment about how realistic the case was. Specifically, 12 of the 20 participants who responded to the open-ended questions indicated that the simulation was realistic or well-done; none of the comments indicated a lack of realism. For example, one participant commented, “I thought this simulation was very similar to my recent audit of a small manufacturing client. I thought that the review notes were insightful, the exercise was realistic, and the simulation, overall, was a great example of a real-life clean-up of accounts receivable/workpaper documentation.”
After the workpapers are returned, participants proceed to a main menu where they can communicate with the client (to obtain explanations and/or additional supporting documentation) and can access the current year working papers with open review notes, excerpts from firm guidance, the accounts receivable audit program, and previously obtained audit evidence. Participants utilize the previously obtained evidence and gather additional evidence to perform the work to address the review notes. Once the review notes have been addressed, participants click on the “finish” tab on the main menu. At this point they access a screen that allows them the option of ending the simulation or going back in case they clicked the finish button in error. If they choose to end simulation, they proceed through a series of manipulation checks and other post-experimental questions before accessing instructions on saving and returning their work.

As alluded to above, participants have the ability to “communicate with the client” by clicking a button on the main menu. When they click on this, they are given more information about the client contacts (e.g., their positions, which customer accounts they oversee) and can choose the appropriate person with whom to communicate. If they choose the incorrect person (e.g., if they ask about an account for which that person is not responsible), they are referred to the appropriate individual.

Once they select an individual, the communication interface asks them to use two to three key words to describe an issue they would like to discuss with the specific client contact. The case is programmed so that many combinations, synonyms, and misspellings of the key words will trigger a response from the client. The pre-testing with auditing students mentioned in section 3.2.2.1 helped in determining these key words, as I was able to extract the words they attempted to use in navigating through the
case. In order to simulate some of the difficulties typically encountered during the information search stage, participants, at times, had to obtain related pieces of evidence from different client personnel (e.g., shipping documents from one person, invoices from another) or were told the client is searching for information and they should return later and inquire again.

### 3.2.2.3 Seeded Errors

The instrument is seeded with the following errors: (1) two invoices are recorded for the incorrect amount; (2) a shipping document does not match the invoice requested (if they follow up on this with client personnel they will learn that the shipping document is missing); (3) two invoices relate to returned merchandise that should have been offset by a credit memo; (4) two write-off memos are not properly authorized; and (5) a credit memo has been issued and recorded after year-end for a return that occurred prior to year-end. Detecting all of these errors involves overcoming nonsampling risk, or the risk of auditor error arising from the possibility that the auditor may: (a) sample the wrong population to test an assertion, (b) fail to detect a misstatement when applying an audit procedure, or (c) misinterpret an audit result (Messier, Glover, and Prawitt 2008). I chose errors of this type because nonsampling risk errors have been shown to be at the heart of many large recent audit failures (Cullinan 2004; Knechel 2007).

---

19 The magnitude (i.e., dollar amount) of each of the errors differs. However, I treat the detection (or non-detection) of each error as equal due to the fact typical audit practice requires any error detected during tests of details to be extrapolated to the population and/or otherwise investigated (Messier, Glover, and Prawitt 2008). Thus, it would not be appropriate to deem any of these errors immaterial. The particular errors used in this study were chosen based on the input and advice of nine practicing and former auditors from three of the Big 4 firms. That is, they were chosen based on my experiences and the experiences of these advising auditors and in such a way that participants with a minimal amount of audit experience should be able to detect them.
The number and type of errors seeded is consistent with previous auditing literature examining auditor detection of nonsampling risk errors in detail testing (Waggoner 1990; Johnstone, Lindsay, and Phillips 2003). Any impact of the number or type of errors on the results of my study would be held constant because the number and type of errors remain constant across conditions and there is no reason to expect that the number or type of errors would differentially impact my independent variables.

### 3.2.3 Independent Variables

The two independent variables (review note frame and review timeliness) are manipulated between participants using a 2 x 2 complete factorial design. I use a between-participants design for the frame condition because it provides the most direct examination of the hypothesized framing effects. Because each review note may require different amounts of information and time to close, issues of comparability would reduce the efficiency gains often present with a within-participants manipulation. In addition, although it is very possible that the same reviewer might frame review notes on the same set of working papers in different ways, a reviewer may have a natural tendency to frame review notes in one particular way. I leave it to future research to examine the effects of multiple frames in one set of working papers.

---

20 Specifically, both Waggoner (1990) and Johnstone, Lindsay, and Phillips (2003) seeded ten deviations or errors into their study which they classified as pertaining to numerical agreement, computation, descriptive agreement, authorization, or reasonableness. They also note that some of the deviations may fall into more than one category. My errors can be classified into these same categories (e.g., the invoices recorded for the incorrect amount would be classified as pertaining to numerical agreement, the incorrect shipping document would be classified as pertaining to descriptive agreement).

21 For instance, I have no reason to believe the number of errors would affect task complexity because finding an error does not require the participant to perform additional work (other than the fact that he or she would have to document the item differently). That is, the amount of work the participant would need to perform in order to fully address the review note is exactly the same regardless of whether they find any of the errors.
There are a total of seven review notes, four of which contain the review note frame manipulation. Each manipulation of a given review note contains the same underlying directive (e.g., “Please check to see that write-offs were properly authorized”); the frame is manipulated as having either a documentation-focused or a conclusion-focused rationale. The rationale in the documentation-framed review note underscores the importance of making sure that there is sufficient, defensible documentation in the workpapers (e.g., “We need to make sure that all of the steps on the audit program have been documented”). In contrast, the rationale in the conclusion-framed review note conveys the importance of the additional work towards ensuring an appropriate conclusion regarding whether the assertion being tested is compliant with GAAP (e.g., “We need to make sure that we arrive at the proper conclusion regarding their authorization assertion”).

Review timeliness is manipulated as delayed or timely. After reading the background information and the instructions, the participant comes to a screen on which they are told that they expect the reviewer to take two days to review the workpapers and they are asked to click on a button to indicate that they are submitting their workpapers to their reviewer. A screen then appears that tells the participant that he or she is waiting for the reviewer to return the workpapers. On this screen, there is a timer that the participant views that counts through the passage of days, with six second intervals representing one day. In the condition in which the review is delayed (timely), the preparer is able to

---

22 The remaining three review notes relate to minor errors in the workpapers that are strictly workpaper presentation issues over which the reviewer is not likely to use much discretion in wording the review note or to include an instructional rationale (e.g., “Please fix the date in the procedures.”). Based on suggestions from former auditors who pre-tested an early version of the instrument, I include review notes of this type to increase the realism of the task.
access the reviewed workpapers after fourteen (two) days. Participants are randomly assigned to one of the four treatment groups.

3.2.4 Pre-Testing of Review Note Frame

I conducted two rounds of pre-testing on the wording of review notes. The first round was performed to ensure that: (1) the review notes are realistic, (2) the manipulation of the focus of the review note rationale does not impact preparers’ understanding of the underlying directive they are being asked to perform (e.g., preparers in the documentation condition should not feel as though they are not actually being asked to perform the work stipulated by the underlying directive), and (3) the rationales provided by documentation and conclusion are viewed by preparers as relatively more documentation-focused and conclusion-focused, respectively. I surveyed 33 auditing students, nine of whom had audit experience, for the first round of pre-testing. I provided these pre-test participants with a brief background description of the audit review process and several sets of review notes (i.e., Review Note Set A, B, C, and D, which correspond with Review Notes 3, 5, 6, and 7 shown on Exhibit 1 of the Appendix). Each set contained a pair of review notes, one with a conclusion rationale and one with a documentation rationale (the underlying directive in each set was identically worded). I then asked the respondents to judge the similarity of the underlying directives of the review note pairs. In addition, I asked them to indicate, for both review notes in each set, the extent to which they felt the reviewer was concerned with (1) the documentation in the workpapers and (2) reaching the appropriate audit conclusions for each review note in

---

23 I chose intervals of six seconds based on informal pre-testing conducted with accounting professors and accounting PhD students with audit experience. The total amount of time the participant waits in the delayed review condition is approximately 84 seconds.
a set. They recorded their responses on 11-point likert scales (0 = not concerned, 10 = very concerned).

As Table 3 shows, results for the four review notes I use in my study indicate that the vast majority of the respondents believed the underlying directives were the same or similar for each set of review notes. Note that only one respondent indicated that he or she viewed the underlying directive for Review Note Set B to be different and only two respondents judged the underlying directives of the note pairing to be different for Review Note Sets A, C, and D. Table 4 presents pre-test participants’ perceptions regarding the relative focus of documentation- and conclusion-framed review notes. Specifically, participants were asked how concerned the reviewer was about the “documentation in the workpapers” and about reaching “the appropriate audit conclusion”. They responded on 11-point likert scales (0 = not concerned, 10 = very concerned). The relative focus was measured by subtracting their assessment of concern for the objective that is not the focus of the frame from their assessment of concern for the objective that is the focus of the frame. In other words, for documentation-framed review notes, the responses for “the appropriate audit conclusion” were subtracted from the responses for “documentation in the workpapers”, with the converse true for

---

24 One pre-test participant, who indicated that the underlying directive was different for all four review note sets still seemed to judge that the procedures should be performed regardless of the focus of the frame. For example, after judging Review Note Set C to have different underlying directives, this participant wrote “1st – stresses making sure cut-off procedures are performed so documentation appears correct. 2nd – stresses making sure procedures are performed to gain comfort over audit evidence” [emphasis added]. The pre-test participant who judged the underlying directives to be different for Review Note Set D appeared to have been confused by the wording used in the review note rather than to have judged any difference in the desire for additional testing to be performed. Specifically, this participant noted “[The documentation-framed review note] wants a basis from A to be tested and checked. [The conclusion-framed review note] wants a basis from ‘X’ to be tested and checked.” It is unclear what this participant meant by “a basis from A” or “a basis from ‘X’”; however, it is important to note that in both cases the participant perceives that the reviewer is asking for “the basis” to be “tested and checked”, not merely documented.
conclusion-framed review notes. One sample t-test comparisons to a mean of 0 are highly significant for each of the review notes utilized in my study. Lastly, six of the experienced pre-test participants were asked to judge the realism of the review notes on 11-point likert scales (0 = not realistic, 10 = very realistic). It appears these participants viewed the review notes as realistic, indicating a mean response of 7.5 (SD = 2.5).

The second pre-test focused on perceptions of reviewers regarding intentional signals the alternative frames may suggest to preparers about the importance and necessity of the additional work requested in the review note (i.e., the underlying directive). Participants were 19 auditors with significant experience reviewing workpapers (mean = 4.92 years reviewing; SD = 3.07). Reviewer pre-test participants were given the four review notes used in my study and were asked to assume they left these notes for a preparer to close. They then responded, on 5-point likert scales, to the following questions for each review note:

a. “Given the wording of this review note, how thoroughly do you expect the staff member to perform the procedure(s) you are requesting?” (1 = not at all thoroughly, 5 = very thoroughly)

b. “Suppose there is an issue or error [in the area being tested by each specific review note] which was not discovered in the additional round of testing. Please indicate your agreement with the following statement: In responding to this review note, I expect that a staff member would perform the procedures requested in such a way that he/she should find the problem.” (1 = strongly disagree, 5 = strongly agree)
c. “Given the wording of this review note, how important do you feel it is that the preparer performs the requested procedures?” (1 = not at all important, 5 = very important)

As Table 5 reports, there are no significant differences between the documentation and conclusion groups for the responses to any of the three questions (all p-values > 0.34). Thus, there is no evidence that reviewers view the alternative frames as sending different signals to preparers regarding the importance and necessity of the additional work requested in the review note, or that they expect the work to be performed any more or less thoroughly with either of the two frames as worded in this study. This suggests that reviewers may not be aware of the effect that their review note phrasing choices have on preparer effort and performance. In other words, any lower levels of effort or performance by the preparer noted in responding to the documentation-framed review notes in this study should not likely be due to an intentional signal from the reviewer to exert less effort while closing the review notes.

3.2.5 Dependent Variables

To measure the amount of supporting evidence examined (H1a, H2a, H3a), I use the cumulative number of relevant evidence items (e.g., invoices, explanations offered by client contacts) accessed by each preparer when closing the review notes. See section 4.4.2 for further discussion of the determination of relevancy of the evidence items. To test H1b, H2b, and H3b, I use the cumulative amount of time the preparers spend addressing review notes (beginning when the workpapers are first returned to preparers until they exit the simulation and begin the post-experimental questions). To measure the performance of the preparer (i.e., the effectiveness of the additional testing performed)
(H1c, H2c, H3c), I count the number of seeded errors the participant discovers and documents. Any documentation indicating the preparer plans to follow-up on the error or seek guidance from the reviewer regarding the error counts as an error discovered.

I classify an error as “discovered” if either (a) the participant discovers, investigates (if necessary), and properly documents the error, or (b) the participant discovers the error but is unsure how to proceed as far as investigating or documenting the error but leaves the review note “open” or documents what he or she found in such a way that the reviewer’s attention would be drawn to the error. In other words, an error is only classified as not “discovered” if the participant documents nothing for the item in question or if the participant documents that everything appears appropriate for the item in question. I use this classification for two reasons. First, the ultimate goal of the study is to make inferences related to the effect of my independent variables on overall audit quality. That is, whether or not a preparer completely understands or makes the correct judgment relating to an item matters less than whether a preparer draws the attention of the reviewer to the item. Second, the coding of the errors (based on the criteria set with the assistance of a former auditor) is completely objective (i.e., there no subjectivity involved with determining that the participant documents either nothing related to the item or that everything is appropriate related to the item).

To examine the use of dysfunctional behaviors (H4a, H4b, and H4c), I focus on the over-documentation of procedures (i.e., instances in which the preparer documents audit procedures without actually performing them).\(^\text{25}\) I code the over-documentation

\(^{25}\) The nature of my instrument allows me to detect (1) over-documentation of audit procedures and (2) whether or not the preparer fails to report material evidence (i.e., the preparer accesses evidence that
independent variable in three ways: (1) a cumulative count of the number of times a preparer documents accessing an item but does not, in fact, access it, (2) binary coding for each participant based on whether he or she engages in any over-documentation (e.g., a participant documents that an invoice was obtained, but never accesses the invoice), and (3) the percentage of over-documentation a particular preparer engages in (calculated as the cumulative count of the number of evidence items over-documented divided by the total number of evidence items accessed by the participant).

In addition to the dependent variables discussed in this section, and the demographic information presented in Table 2, participants responded to several other post-experimental questions. These questions relate to affect, the way in which the preparer attributed the timeliness of the review, and the preparer’s perception of the reviewer’s concern for workpaper documentation, audit conclusions, etc. These measures are discussed in section 4.2.

indicates there is an error but does not document it). I can easily, and objectively, determine whether a participant over-documents because I have built tracking variables into Authorware in such a way that a data file tracks each evidence item that a participant accesses. Due to the more subjective nature of the second type of dysfunctional behavior (i.e., I can determine whether a participant accessed an item that should alert him or her to the error, but not whether he or she actually read the item or understood the information provided), I focus on over-documentation.
CHAPTER 4: RESULTS

4.1 Introduction

In the previous chapter I developed hypotheses, discussed the design of the study, and presented results of pre-testing of the experimental instrument. In this chapter, I provide results of manipulation checks, related post-experimental questions, and hypotheses testing. In order to test my hypotheses, I analyze participants’ responses using a 2x2 analysis of variance (ANOVA) with the frame of the review note and the timeliness of the review as my independent variables. The test examines the effect of review note frame and reviewer delay on (a) the number of supporting evidence items examined (H1a, H2a, and H3a), (b) the time spent closing review notes (H1b, H2b, and H3b), (c) the number of errors identified (H1c, H2c, and H3c), and (d) the amount of over-documentation of procedures (H4a, H4b, and H4c). All tests will be one-tailed because the hypotheses are directional.

4.2 Manipulation Checks and Post-Experimental Questions

4.2.1 Review Timeliness

Participants were randomly assigned to one of two review timeliness conditions: (1) a timely review which matched the expectation set of two days and (2) a delayed review (14 days) which exceeded the expectation of two days. After completing the simulation, I asked participants, “In the case you just completed, how long did the manager take to return the workpapers to you?” (2 days/14 days). All 69 participants correctly answered this question. Because I set an expectation of two days for the review in both conditions, I also asked participants: “In the case you just completed, how long
did you expect the audit manager to take to return the workpapers to you (i.e., what was the anticipated turnaround time for the review)?” (2 days / 14 days). Six participants incorrectly answered this question. All of the participants who did not answer correctly were in the timely review condition and removing them from the analysis does not significantly affect results of hypotheses testing or any inferences drawn from those results.

The timeliness of the review is expected to impact performance by causing the participants to feel negative affect and/or impact the preparers’ perception of the priority of the area to the reviewer. In addition, Sheldon, Thomas-Hunt, and Proell (2006) find that time delays may undermine the influence or, in the case of a high-status delayer, the authority of the delayer. Therefore, I collect data on items in which there is a directional expectation related to the timeliness manipulation in order to determine how participants experienced the reviewer delay. Participants responded to the following items on 11-point likert scales:26

a. “In the case you just completed, do you feel the length of time the audit manager had the workpapers before returning them to you was long or short?”
   (1 = very short, 11 = very long)

b. “How much of a priority do you feel these workpapers were to the manager?”
   (1 = low priority, 11 = high priority)

c. “To what extent did the length of review time affect your perceptions of the manager’s authority over you/the audit (i.e., their influence/power as an

26 Because the instrument was lengthy and comprehensive, there may be some general baseline frustration with the instrument that would be randomized across conditions. Therefore, it was important to design questions to measure negative affect specifically related to review timelines rather than use the specific wording of questions used in previous studies involving negative affect.
authority figure on the audit?” (1 = strongly decreased view of authority, 6 =
did not impact, 11 = strongly increased view of authority)

d. “To what extent were you frustrated by the length of his review time?” (1 =
not at all frustrated, 11 = very frustrated)

e. “To what extent did you feel angry due to the length of his review time?” (1 =
not at all angry, 11 = very angry)

Also, as discussed in chapter 2 and during the hypotheses development in chapter
3, the impact of time delay may depend on the way in which the preparer attributes the
delay. Although the attribution may help to determine the way in which the preparer
reacts to the timeliness of the review, I do not necessarily expect different attributions
based on the timeliness condition. That is, I have no reason to expect that a preparer will
attribute a timely review in one way and a delayed review in a different way (e.g.,
attribute a delayed review to self and a timely review to the manager). I measure the
following items to determine how participants attribute the timeliness of the review on
11-point likert scales:

f. “How much of a control do you believe the audit manager had over the length
   of time he took to review the workpapers?” (1 = no control, 11 = significant
   control)

g. “How much do you feel that the length of time the manager took to review the
   workpapers was due to factors related to you, personally (as opposed to
   factors related to the manager and/or audit environment)?” (1 = not at all
   related to me, 11 = very much related to me)
h. “How much do you feel that the length of time the manager took to review the workpapers was due to factors related to the manager and/or audit environment (e.g., manager’s time management skills, pressures associated with manager’s workload or the time of the year, manager’s prioritization of this workpaper)?” (1 = not at all related to manager/environment, 11 = very much related to manager/environment)

Table 6 presents analyses of participants’ responses to these post-experimental questions. Panel A presents independent t-test results for questions a. through e. above using one-tailed p-values as there are directional expectations related to the timeliness manipulation for these questions. As anticipated, participants in the delayed review condition perceived the length of time the manager took to review the workpapers (mean = 10.23, SD = 1.14) as significantly longer than participants in the timely review condition (mean = 5.18, SD = 2.24; p < 0.001). Consistent with expectations based on the literature review, participants in the delayed review condition reported significantly higher levels of frustration and anger (means = 8.17 and 5.89, SD = 1.92 and 2.73, respectively) than the timely review condition (means = 2.68 and 1.88, SD = 2.54 and 1.70, respectively; both p’s < 0.001). Also consistent with previous literature, delayed review participants reported that they perceived the workpapers to be of significantly lower priority to the manager (mean = 2.40, SD = 2.00) than their counterparts in the timely review condition (mean = 6.00, SD = 2.74; p < 0.001), and the delay appears to have had a significant effect on participants’ perceptions of the manager’s authority (p = 0.041). Timely review participants reported higher perceptions of the manager’s authority (mean = 6.65, SD = 1.97) than delayed participants (mean = 5.69, SD = 2.52).
As discussed above, although the attribution a preparer makes may impact the way in which he or she experiences and/or responds to the delay, there is no reason to expect that participants in the timely condition will attribute the timeliness of the review differently than participants in the delayed condition. Panel B of Table 6 reports results of independent t-tests for these questions in which there is no directional expectation for the differences between conditions. As such, the p-values for these tests are two-tailed. Participants appear to perceive that the manager has some level of control over the delay as the mean response to question f. for both the timely condition (mean = 7.71, SD = 2.50) and the delayed condition (mean = 7.40, SD = 2.13) is significantly greater than the mid-point of the 11-point scale (both p’s < 0.001 for t-test comparison to mid-point of 6). The responses for both conditions are not significantly different from each other (p = 0.585). Participants in neither group appear to attribute the delay to themselves as the means for both the timely group (mean = 3.68, SD = 2.37) and the delayed group (mean = 3.63, SD = 2.00) are well below the mid-point of the response scale (both p’s < 0.001 for t-test comparison to mid-point of 6). Again, the responses for both conditions are not significantly different from one another (p = 0.928). Both timeliness conditions do seem to attribute the delay to the manager/environment, as mean responses for both the timely condition (mean = 8.62, SD = 1.83) and the delayed condition (mean = 8.37, SD = 1.80) are above the mid-point of the scale for question h. (both p’s < 0.001 for t-test comparison to mid-point of 6). The responses for this question for timely and delayed participants are not significantly different from each other (p = 0.575). In summary, it appears that participants in both conditions feel that the timeliness of the review is a result of some combination of the manager and the audit environment. Based on the high
scores reported for question h., participants appear to view the manager as having a considerable level of control over the delay, suggesting that they are attributing much of the responsibility for the delay to the reviewer.

4.2.2 Review Note Frame

O’Keefe (2003) asserts that, when examining the impact of a message variation on an outcome, there is no need to check the manipulation of the message by assessing participant perceptions. That is, he argues that a manipulated message either does or does not contain the manipulation and that anything reported as a manipulation check is really an assessment of a potential mediating variable. According to O’Keefe (2003, p. 257), “The investigator will naturally want to be careful in creating the experimental messages, but the adequacy of the manipulation of the message property is not appropriately assessed by inquiring about participant perceptions of the message.” Following this advice, great care was taken in creating the review note manipulations (i.e., they were created with the input of former and practicing auditors) and the review notes were extensively pre-tested as described in section 3.2.4. Thus, there are no manipulation checks, per se, for review note frame. However, I did ask post-experimental questions related to the frame of the review note as described below.

As discussed in section 3.1.2.1, two important objectives of the review process are: (1) reaching an appropriate conclusion regarding conformance with generally accepted accounting principles (GAAP) and (2) ensuring that the documentation in the workpapers is defensible and will withstand ex post scrutiny (Rich, Solomon, and Trotman 1997a). Participants in this study are audit staff and seniors who have been primed during their accounting education and careers to be aware of, and concerned with,
both of these objectives. In other words, I expect all of my participants to be concerned with both conclusion and documentation concerns. The manipulation is intended to have an impact on the extent to which one concern is relatively more or less salient in the mind of the participant.

Therefore, I measured the extent to which the participant perceived that the reviewer was concerned with each of the two primary objectives. In addition, I expect that the conclusion-frame (documentation-frame) might increase (decrease) the extent to which the preparer perceives that the manager believes there are undiscovered issues or errors in client’s accounts. Therefore, I measure the following items on 11-point likert scales:

a. “In general, how concerned was the reviewer about the documentation in the workpapers?” (1 = not concerned, 11 = very concerned)

b. “In general, how concerned was the reviewer about reaching the appropriate audit conclusions?” (1 = not concerned, 11 = very concerned)

c. “In general, how concerned do you believe the audit manager was that there may be undiscovered issues or errors in Sprandel Inc’s Accounts Receivable?” (1 = not concerned, 11 = very concerned)

Table 7 presents analyses of participants’ responses to these post-experimental questions and analysis of the perception of the relative concern that the manager has for reaching the appropriate audit conclusions vis-à-vis concern for the documentation in the workpapers (the participant’s response to question b. above minus the response to question a. above). All p-values are one-tailed due to directional expectations related to the manipulation. Preparers in the conclusion-frame condition perceive a significantly
lower level of concern by the reviewer for the documentation in the workpapers (mean = 6.47, SD = 2.51) than preparers in the documentation-frame condition (mean = 7.84, SD = 2.42; p = 0.013). However, while differences are in the expected directions, they are not significant for participants’ perceptions of the manager’s concern for reaching the appropriate audit conclusions (p = 0.202) or that there may be undiscovered issues or errors in the client’s Accounts Receivable (p = 0.187). Means for both questions, however, are in the predicted direction. Conclusion-framed review note respondents reported a mean perceived concern for the appropriate audit conclusions of 8.44 (SD = 2.03) and a mean perceived concern that there may be undiscovered issues or errors of 7.53 (SD = 2.08). The documentation-frame condition reported means of 7.95 and 7.03 (SD = 2.82 and 2.59) for these two questions, respectively. The lack of significance may be because the main purpose of an audit is to reach an appropriate conclusion as to whether the financial statements conform to GAAP. Auditors learn the importance of this concept from the beginning of their undergraduate accounting education. Thus, it is probable that all participants feel that the manager is likely to be concerned about this objective and about undiscovered issues or errors in the account balance being tested.

As outlined in section 2.3.2 and 3.1.2.1, an emphasis frame works by influencing the relative salience of potential concerns. Thus, it is important to examine a relative measure of the participant’s perception of the manager’s concern for reaching the appropriate conclusion vis-à-vis concern for the documentation in the workpapers. The difference between the perception of the manager’s concern for reaching an appropriate conclusion (item b.) and ensuring the documentation is defensible (item a.), is significant for review note frame (means = 1.97 and 0.11, SD = 2.73 and 2.96 for the conclusion and
documentation frames, respectively; \( p = 0.009 \). Thus, participants’ perceptions of the relative importance of these two primary review note objectives differs significantly in such a way that conclusion-framed respondents perceive the manager as relatively more concerned with reaching the appropriate audit conclusion than documentation-frame respondents.

4.3 Testing of Statistical Assumptions

There are three assumptions that must be met in order to use ANOVA to test the hypotheses presented in chapter 3: (1) independence of the dependent variable between treatment groups, (2) normal distribution of treatment populations, and (3) homogeneity of variance between treatment populations (Keppel 1991; Gardner 2001). Participants were randomly assigned to the four experimental conditions. That is, the assignment of a participant was not dependent upon the assignment of any other participant. Thus the first assumption is met (Keppel 1991; Gardner 2001). I use the Kolmogorov-Smirnov test of normality to examine the second assumption. Non-tabulated results specify that all three dependent variables appear to be normally distributed (all two-tailed \( p \)’s > .10), and thus the second assumption is met. I use Levene’s test of homogeneity of variances to examine whether the error variance of the dependent variables is equal across groups. Non-tabulated results indicate that the variance among treatment populations is homogenous (all two-tailed \( p \)’s > .10). Therefore, I conclude that all three assumptions are met, and thus all of the hypotheses testing will be performed using ANOVA.
4.4 Hypotheses Testing

4.4.1 Introduction

This section presents the results of my hypotheses testing. The tests explore the impact of review timeliness and review note frame on audit workpaper preparer effort and performance while closing review notes. The main hypotheses developed in chapter 3 predict relationships between the two dichotomous independent variables and three dependent measures of effort and performance (number of evidence items examined, time spent closing review notes, and number of errors found). I also examine the influence of the two independent variables on the extent to which the preparer engages in dysfunctional behavior in the form of over-documentation of procedures. Individual hypothesis testing is conducted using a 2x2 ANOVA model and contrast tests are conducted where applicable.27 One-tailed tests are used to analyze results when there are directional expectations, and two-tailed tests are used otherwise.

4.4.2 Number of Evidence Items Examined

As a measure of audit effort, I examine the cumulative number of relevant evidence items examined. I expect review timeliness and review note frame to affect the number of items examined (H1a, H2a, and H3a). I have built tracking variables into the computerized case to determine which evidence items each participant accesses. Relevant evidence items are those items that are necessary to fully address a review note.28 For example, the review note on subsequent receipts testing requires the

27 Given that there are multiple dependent variables (two measures of effort and one measure of performance), multivariate analysis of variance was performed and results indicate significance of the model (p < 0.001).
participant to: (1) examine invoices, (2) examine shipping documents, and (3) inquire of
the client whether there has been communication with customers. If they perform all
three of these steps for a specific selection, they have looked at three evidence items for
that selection. If, for instance, they examine the invoices and shipping documents but do
not access the information they would have been provided if they had inquired about
correspondence with the customers, then they have looked at two evidence items for that
selection.

Hypotheses 1a, 2a, and 3a predict a main effect of review timeliness, a main
effect of review note frame, and an interactive effect of review note frame and review
timeliness and are stated as follows:

**H1a**: Preparers will examine more supporting evidence when closing conclusion-frame
review notes than when closing documentation-framed review notes.

**H2a**: Preparers will examine less supporting evidence when there is a reviewer delay
than when there is no reviewer delay.

**H3a**: Review timeliness and review note frame will interact such that the positive effect
conclusion (relative to documentation) frame on the amount of supporting
evidence preparers examine is greater when the review is timely that when the
review is delayed.

---

I have not specifically built “irrelevant” or “unnecessary” evidence items into the case. However, there
are instances in which participants may access support (e.g., account detail) that adds to the realism of the
task and may help them to obtain relevant evidence items from the client, though that support is not entirely
necessary. If they obtain the relevant evidence items without accessing this support, then they will be
considered as having obtained all of the relevant evidence items for that review note. Determination of
relevant evidence items is not subjective, as the review notes either specifically state what evidence items
the preparer should access or they refer the preparer to the audit program/firm guidance which specifically
states which items the preparer should access. Furthermore, it is possible for the preparer to find an error
without accessing all pieces of relevant evidence items. However, the preparer should access all evidence
items in order to fully comply with the review note.
Table 8 presents the results of the 2x2 ANOVA (review timeliness by review note frame) using total evidence items examined as the dependent variable. Reported significance levels are one-tailed because directional hypotheses are posited. Table 8, Panel A shows a marginally significant main effect for review note frame (p = 0.066). Preparers closing conclusion-framed review notes examine an average of 28.25 items (SD = 10.33) and preparers closing documentation-framed review notes examine an average of 24.49 items (SD = 10.18; Table 8, Panel B) showing marginal support for H1a. Table 8, Panel A also reports that preparers receiving a timely review examine significantly more evidence items than those subject to a delayed review (p = 0.049). Specifically, participants in the timely review condition examine an average of 28.26 evidence items (SD = 10.02) and participants in the delayed review condition examine an average of 24.26 items (SD = 10.42; Table 8, Panel B). Thus, H2a is supported. The interaction of review note frame and the timeliness of the review is not significant as shown on Table 8, Panel A (p = 0.193). While this result does not support H3a, results of contrast tests point to an interesting finding relating to the combined effect of a timely review and conclusion-framed review notes. Specifically, participants in Cell 1 examine significantly more items (mean = 31.38) than each of the other cells (means for Cells 2 through 4 = 25.50, 25.13, and 23.53, respectively; p’s < 0.05), while the others are not significantly different from one another (p’s > 0.27). Thus, it appears that the results relating to number of evidence items examined are largely driven by the superior performance of participants in Cell 1 (i.e., the combined impact of conclusion-framed

29 As displayed on Table 1, cells are defined as follows: Cell 1 = timely review, conclusion frame; Cell 2 = timely review, documentation frame; Cell 3 = delayed review, conclusion frame; Cell 4 = delayed review, documentation frame.
review notes with a timely review). That is, the frame matters most when the review is timely.

4.4.3 Time Spent Closing Review Notes

As another measure of effort, I examine the amount of time preparers spend closing review notes. Due to the extensive nature of my instrument and the amount of time required to complete it (mean = 83.60 minutes, standard deviation = 51.75 minutes), it appears that some participants did not complete the entire simulation without pausing. Therefore, I perform an analysis of the time stamps recorded in the data file of each participant to adjust for the amount of time a participant may not have been actively working on the instrument.\footnote{The program stamps the time whenever a participant clicks on something within the authorware program. When a participant works within the workpapers (i.e., documenting evidence or conclusions on the excel spreadsheet which comprise the workpapers), there would be no time stamping of the participant’s activity. This time stamping is not apparent to the participant (i.e., there is no clock visible to the participant within the simulation itself).} For pauses greater than 20 minutes, I subtract the amount of time greater than 20 minutes for each participant.\footnote{I exclude any time beyond 20 minutes during which the participant is paused from the total time measure (i.e., any pauses greater than 20 minutes are counted as 20 minutes). I analyzed the data using cut-offs of 5, 10, 15, 20, 25, and 30 minutes. Neither the significance of the results nor the inferences drawn from them change based on the use of these various cut-off points. I present the primary results using a 20 minute cut-off due to the high likelihood that pauses over 20 minutes are due to the participant taking a short break (e.g., getting coffee, answering the phone), rather than the participant working on the simulation (e.g., processing information, documenting in the workpapers). A total of 8 participants had pauses greater than 20 minutes. However, due to the arbitrary nature of the 20 minute cut-off, I perform a sensitivity analysis using a 10 minute cut-off and the rank-order of time as dependent variables (see Tables 10 and 11).} I then perform a Q Plot of the data and note one very obvious outlier (total adjusted time spent closing review notes = 196.43 minutes; see Figure 3), which I replace with the next highest value for time present in the sample.\footnote{See section 4.4.7 for a discussion of the results for all analyses if this outlier is omitted. In addition, I examined (1) the number of participants with delays over these various cut-off points per cell and (2) the total number of minutes over these various cut-off points per cell. There were no significant differences} I use these adjusted total times spent closing review notes to test
H1b, H2b, and H3b. Hypotheses 1b, 2b, and 3b predict a main effect of review timeliness, a main effect of review note frame, and an interactive effect of review note frame and review timeliness, respectively, and are stated as follows:

**H1b:** Preparers will spend more time closing conclusion-framed review notes than closing documentation-framed review notes.

**H2b:** Preparers will spend less time closing review notes when there is a reviewer delay than when there is no reviewer delay.

**H3b:** Review timeliness and review note frame will interact such that the positive effect conclusion (relative to documentation) frame on the amount of time preparers spend closing review notes is greater when the review is timely than when the review is delayed.

Table 9 presents the results of the 2x2 ANOVA (review timeliness by review note frame) using the total time the preparer spends closing the review notes (adjusted for pauses greater than 20 minutes) as the dependent variable. Table 9, Panel A shows that H1b is not supported, as the main effect of review note frame is not significant (p = 0.256). The means, however, are in the predicted direction. Preparers who receive conclusion-framed review notes spend an average of 80.18 minutes (SD = 34.34) while preparers who receive documentation-framed review notes spend an average of 74.50 (SD = 36.44; Table 9, Panel B). Results indicate that preparers receiving a timely review do spend more time closing review notes than those receiving a delayed review (p = 0.087), providing support for H2b (Table 9, Figure A). Participants spent an average of 82.82 among the cells for these measures indicating that the occurrence of participants taking such pauses was randomly distributed among cells and, thus, is unlikely to impact the results of this study in a systematic way.
minutes (SD = 33.40) closing the review notes in the timely condition in comparison to 71.61 minutes (SD = 36.77) in the delayed condition. Table 9, Panel A indicates that review note frame and review timeliness do not interact in a significant way to influence the amount of time spent by the preparer (p = 0.167), indicating a lack of support for H3b. However, again, Cell 1 appears the most different of the four cells. Contrast tests in Panel C of Table 9 show that participants in Cell 1 spend significantly more time closing review notes (mean = 90.20) than those in both Cells 3 and 4 (means = 70.16 and 72.84, respectively; p’s < 0.10). Again, Cell 1 appears to be the primary driver of these results. The combination of conclusion-framed review notes and a timely review appears to have the greatest effect on the amount of time preparers spend closing review notes.

I performed two additional robustness tests to ensure that the results reported in the above paragraph are not driven by the way in which I analyzed the participants’ time stamps. For the first robustness test, I perform the same analysis described above adjusting for pauses greater than ten minutes instead of five minutes. Once again, I replaced an outlier (total adjusted time spent closing review notes = 227.75) with the next highest value (total adjusted time spent closing review notes = 152.30). Table 10 presents the results of an ANOVA using the total time the preparer spends closing the review notes (adjusted for pauses greater than ten minutes) as the dependent variable. As can be seen, the results are qualitatively similar to those reported in Table 9. Again, Hypothesis 2b is supported (Table 10, Panel A), with preparers spending marginally more time closing review notes when the review is timely than when it is delayed (p = 0.059). There is no significant main effect for review note frame (p = 0.314) and there is no significant interaction (p = 0.149). Therefore, H1b and H3b are not supported. As with
the main analysis, contrast tests (Table 10, Panel C) show that participants in Cell 1 spend significantly more time (mean = 86.36 minutes) closing review notes than those in Cells 3 and 4 (means = 65.88 and 70.26, respectively; p’s = 0.038 and 0.073, respectively).

For the second robustness test, I use the rank of the time spent closing review notes as the dependent variables (without adjusting for pauses). Rank involves recoding the data into their rank ordering with the shortest time given a value of one and the longest time (based on my sample) given a value of 69. Table 11 displays the results of this analysis and shows the inferences drawn are consistent with both Tables 9 and 10. Table 11, Panel A shows (1) a marginally significant main effect of review timeliness (p = 0.079), (2) no significant main effect of review note frame (p = 0.133), and (3) no significant interaction (p = 0.304). Thus, again, only H2b is supported. The results of the contrast tests presented in Table 11, Panel C are consistent with Panel C of Tables 9 and 11. Participants in Cell 1 spend significantly more time (mean = 42.63) closing review notes than both Cells 3 and 4 (means for Cells 3 and 4 = 33.25 and 30.32, respectively; p’s = 0.095 and 0.037, respectively). In summary, all three measures of time spent closing review notes demonstrate support for H2b (review timeliness) and directionally consistent, but insignificant, results for both H1b (review note frame) and H3b (the interaction of review timeliness and review note frame), indicating a lack of support for these two hypotheses.

4.4.4 Number of Errors Found

To measure the performance of the preparer (i.e., the accuracy of the results of the additional testing performed), I count the number of seeded errors the participant
discovers and documents (H1c, H2c, H3c). As discussed in Section 3.2.5, I count any documentation indicating the preparer plans to follow-up on the error or seek guidance from the reviewer regarding the error as an error discovered. Hypotheses 1c, 2c, and 3c predict a main effect of review note frame, a main effect of review timeliness, and an interactive effect of the two variables, respectively, and are stated as follows:

H1c: Preparers will identify more errors when closing conclusion-framed review notes than when closing documentation-framed review notes.

H2c: Preparers will identify fewer errors when there is a reviewer delay than when there is no reviewer delay.

H3c: Review timeliness and review note frame will interact such that the effect of note frame on the number of errors preparers identify is greater when the review is timely than when the review is delayed.

Table 12 presents the results of the 2x2 ANOVA (review timeliness by review note frame) using the total errors found as the dependent variable. Panel A of Table 12 shows no significant effect for review note frame (p = 0.279) although the means are in the predicted direction, with preparers who receive conclusion-framed review notes finding a mean of 3.00 errors (SD = 1.50) and those who receive documentation-framed review notes finding a mean of 2.78 errors (SD = 1.58). Thus, the results do not support H1c. Table 12, Panel A indicates a main effect for review timeliness (p = 0.067) which supports H2c. Reviewers responding to a timely review find an average of 3.15 errors (SD = 1.50) in comparison to delayed preparers who find an average of 2.63 errors (SD = 1.56). However, care must be taken in drawing inferences regarding this main effect due to the fact that the mean of Cell 2 (the timely/documentation condition) is nearly identical
to the mean of Cell 4 (the delayed/documentation condition) (means = 2.78 and 2.79, respectively). Table 12, Panel A does show support for H3c as the interaction is marginally significant at $p = 0.063$. The interaction appears to be driven by the superior performance exhibited by participants in Cell 1, with Panel C contrast tests showing that the mean number of errors found in this cell (3.56 errors, SD = 1.21) is significantly higher than the means of the other three cells (2.78, 2.44, and 2.79, respectively, for Cells 2, 3, and 4).

In light of the results of these contrast tests and the form this interaction takes, it is prudent to reassess the support for H2c. It appears that the significant main effect of reviewer delay does not represent a true main effect. That is, there is no effect of delay on documentation-framed review notes (i.e., Cells 2 and 4 are virtually identical). The ANOVA results indicating a significant effect of reviewer delay are therefore attributable to the difference between Cells 1 and 3. In summary, with respect to errors found by the preparer it appears that only H3c is supported.33

33 The dependent variables used to test H1, H2, and H3 (i.e., number of supporting evidence items, amount of time spent closing review notes, and number of errors identified) are highly correlated ($p < 0.001$ for all three simple correlations). The high correlation is not surprising because greater levels of effort are expected to lead to better performance. Ideally, a mediation analysis would be in order to consider the path through which the independent variables impact the number of errors found. Recall, however, that the majority of my results are driven by Cell 1 (i.e., the combination of conclusion-framed review notes and review timeliness). Unfortunately, there does not appear to be an accepted method of testing for mediation of an interaction of two dichotomous variables. Therefore, I explore mediation analysis using each independent variable separately. For mediation to exist, the effect of the independent variable on the final dependent variable (in my case, errors found) must be significant unless (1) the measurement of the dependent variable is far removed in time from the measurement of the independent variable and (2) the mediated and direct effects have different signs (Taylor, MacKinnon, and Tein 2008). There is no main effect of review note frame on the number of errors found and the previously mentioned two criteria are not met. Thus, I do not perform a mediation analysis using review note frame as the independent variable. I do, however, perform a multiple mediation analysis using review timeliness as the independent variable following Baron and Kenny (1986). Results from this analysis suggest that the effect of reviewer delay on the number of errors found is fully mediated by the two effort measures ($p$-value on the review timeliness coefficient after controlling for the effects of the two mediators is 0.329). In addition, the multiple mediator model suggests that evidence items examined is a stronger mediator than time spent as the
4.4.5 Over-Documentation of Procedures

As a measure of dysfunctional behavior, I examine the extent to which preparers document that they have examined evidence items which they have not, in fact, examined. The instrument records each evidence item the participant accesses and, thus, allows me to compare what items preparers document they examine with the items they actually examine. I examine over-documentation using three measures: (1) a cumulative count of the number of times a preparer documents accessing an item but does not, in fact, access it, (2) binary coding for each participant based on whether or not he or she engages in any over-documentation (i.e., participants who over-document at least one item are coded as 1; those who do not over-document at all are coded as 0), and (3) the percentage of over-documentation in which a particular preparer engages (calculated as the cumulative count of the number of evidence items over-documentated divided by the total number of evidence items accessed by the participant). I use these three measures to test H4a, H4b, and H4c, which are stated as follows:

**H4a:** Preparers receiving a timely review will engage in less dysfunctional behavior while closing review notes than preparers receiving a delayed review.

**H4b:** Preparers receiving conclusion-framed review notes will engage in less dysfunctional behavior while closing review notes than preparers receiving documentation-framed review notes.

---

Coefficient on the former has a p-value of 0.001 and the latter has a p-value of 0.081. I footnote these results, rather than include them in the body of the text, because there is limited knowledge gained due to the lack of an appropriate method for investigating mediation of the interaction.
**H4c:** Review timeliness and review note frame will interact such that the positive effect of note frame on the propensity to engage in dysfunctional behavior is greater when the review is timely than when the review is delayed.

Table 13 presents the results for the first two measures of over-documentation. Panel A shows that, while results are in the expected direction, participants’ cumulative over-documentation for conclusion-framed review notes is not significantly different than for documentation-framed review notes (mean = 2.44 and 2.62, respectively; p = 0.439). There is a larger, though again insignificant, difference in the mean cumulative amount of over-documentation between preparers in the timely and delayed review conditions (mean = 1.91 and 3.14, respectively; p = 0.131). In addition, the interaction is not significant (p = 0.370). Panels C and D show similar results using the proportion of participants who over-document as the dependent variable. Again, results are in the expected direction, though not significant, when looking at the main effect of review note frame with 31% of participants over-documenting in the conclusion-frame condition and 35% of participants over-documenting in the documentation-frame condition. Also consistent with the first measure, there is a larger, though again insignificant, difference in the propensity to over-document between preparers in the delayed review condition (40% of participants over-document) and those in the timely review condition (26%). In addition, the interaction is not significant (p = 0.469).

Table 14 presents the results using the third measure of over-documentation as the dependent variable: the cumulative amount of evidence items over-documentated divided by the total evidence items examined. Results for this analysis do show support for H4a. Preparers receiving a timely review over-document significantly fewer evidence
items as a proportion of total evidence items examined ($p = 0.043$). Consistent with the first two measures of over-documentation, H4b and H4c are not supported ($p = 0.316$ and 0.158, respectively). Contrast tests show that preparers receiving a timely review in both frame conditions (Cells 1 and 2) over-document significantly fewer evidence items as a percentage of total evidence items examined than preparers receiving a delayed review and conclusion-framed review notes (Cell 3) ($p$’s $= 0.031$ and 0.060, respectively). Thus, the results indicate mixed support for H4a and do not support H4b and H4c.

**4.4.6 Controlling for Demographic and Other Post-Experimental Variables**

All analyses used to perform hypotheses testing were also performed using the demographic and other (non-manipulation check) post-experimental variables presented in Tables 2, 6, and 7 as covariates. Controlling for these variables does not significantly affect the results presented or change any of the inferences drawn.

**4.4.7 Results of Analyses with Time Outlier Omitted**

As discussed in Section 4.4.3 and shown in Figure 3, there was one outlying observation that had an extreme value for the time variable. Upon examination of the time stamps for this individual, I noted several long pauses in the simulation. The two longest pauses are for lengths of time of over 131 minutes and over 22 minutes. In case these pauses caused this individual to experience the manipulations differently (particularly the time delay manipulation), I also perform all analyses omitting this individual. Results using the number of evidence items examined as the dependent variable do not significantly change (i.e., the timeliness main effect is significant and the main effect of review note frame is marginally significant). Using time spent as the dependent variable (the first measure of time spent presented in Table 9), provides more
significant results for the main effect of timeliness (p = 0.034, as opposed to 0.067 on Table 9) and marginally significant results for the interaction (p = 0.078, as opposed to 0.146 on Table 9).\textsuperscript{34} Finally, deleting this outlier causes the interaction of the independent variables on the number of errors noted by the participant to increase in significance (p = 0.049, as opposed to 0.063 on Table 12). The main effect of delay for this dependent variable is still at a marginally significant level. Omitting this outlier does not affect the results for the dysfunctional behavior dependent variable measures. Thus, deleting this outlier appears to only improve the results of hypotheses testing.

4.5 Summary of Hypotheses Testing

The results of this study indicate that both review timeliness and review note frame affect levels of effort and performance of preparers closing review notes. A timely review leads to significantly higher levels of effort (as measured by total of number of evidence items examined and total time spent closing review notes) and performance (as measured by the number of errors identified). Conclusion-framed review notes lead to higher levels of effort (as measured by total number of evidence items examined).

Review note frame and review timeliness interact in a significant way to impact the number of evidence items examined. With respect to the number of errors found, both the main effect of review timeliness and the interactive effect of timeliness and review note frame appear to be primarily driven by the superior performance of the timely review/conclusion-framed review note condition. That is, participants in this condition

\textsuperscript{34} The results for the measure of time spent in which pauses over 10 minutes are deleted (i.e., the analysis presented in Table 10) change in a similar way to those discussed in this section for the measure of time spent in which pauses over 20 minutes are deleted. Specifically, deleting the outlier causes the interaction to become marginally significant using this measure. However, the interaction is not significant using the measure of the rank of time as the dependent variable (i.e., the measure of time analyzed at Table 11).
find significantly more errors than participants in the other three conditions. In addition, the results provide some mixed evidence to suggest that a delayed review may lead to a greater propensity for preparers to over-document evidence than a timely review.
CHAPTER 5: CONCLUSIONS, LIMITATIONS, AND IMPLICATIONS

5.1 Introduction

This study examines the effects of review note frame and review timeliness on preparers’ effort and performance while closing review notes (i.e., responding to comments written by the reviewer which require follow-up by the preparer). Experienced auditors completed an experimental audit simulation which asked them to assume the role of an audit workpaper preparer. They then submitted workpapers and later received review notes back from their audit manager, which they needed to close. The review notes were manipulated so that the rationale for requiring the additional work either focused on the sufficiency of documentation or on the appropriateness of conclusions drawn from procedures performed on the audit. The specific task (i.e., the underlying directive) requested by the reviewer within each type of review note were identical. The timeliness of the review was manipulated as having been timely (expecting and taking two days) or delayed (expecting two days and taking fourteen days). The following sections offer conclusions, limitations, and implications for this study.

5.2 Conclusions

Prior research has recognized the importance of the audit review process and has examined such factors as whether a reviewer detects an error in the preparers’ work, the amount of effort a reviewer expends, the type of review note a reviewer leaves, and whether a reviewer anticipates stylization attempts by a preparer (Ballou 2001; Gibbins and Trotman 2002; Tan and Trotman 2003). While a substantial body of research on the
audit review process is developing, prior research has not investigated the phase of the audit in which preparers respond to and close review notes. Thus, the extant literature cannot speak to factors that impact whether or not an issue eventually gets resolved once a reviewer leaves a review note. This study is the first to examine the review note closing process and to attempt to understand factors that affect the amount of supporting documentation examined, the time spent addressing the review notes, and the results of any additional work performed by the preparer.

As noted in previous literature, the procedures performed by junior staff members serve as the foundation of the audit opinion and, thus, errors or problems in their work could potentially lead to significant audit failures (Willet and Page 1996; Herrbach 2005). In addition, there is survey and anecdotal evidence to suggest that audit staff and seniors engage in dysfunctional behavior (such as failing to report material evidence, signing off on an audit step without actually completing it, and biasing the sample selection) that may threaten the quality of the audit (Pierce and Sweeney 2006; Donelly, Quirin, and O’Brien 2003; Rich, Solomon, and Trotman 1997a; Otley and Pierce 1996). This study is the first to provide experimental evidence on the tendency for auditors to engage in such behavior (specifically, to over-document procedures performed in the workpapers).

The results suggest that there are reviewer-related contextual factors that can affect the effort and performance of preparers closing review notes. Providing a timely review of the workpapers can enhance the levels of effort and performance of preparers while closing review notes. This study suggests that preparers receiving timely reviews examine significantly more evidence items and spend more time closing review notes than preparers receiving delayed reviews. In addition, this study finds that conclusion-
framed review notes lead to higher effort levels (as measured by the number of evidence items examined) than documentation-framed review notes. Also, review timeliness and the frame of the review note interact in such a way that the frame of the review note affects the number of errors found more when there is a timely review than when the review is delayed. In addition, contrast tests on the other dependent variables (number of evidence items examined and time spent closing review notes) indicate that preparers receiving a timely review with conclusion-framed review notes generally exhibit superior levels of effort and performance than that exhibited by preparers in each of the other conditions. That is, the frame appears to matter most when the review is timely (i.e., the combination of conclusion-framed review notes and a timely review generally leads to the best outcome). This finding corroborates emphasis framing literature that suggests the effect of a message frame may be moderated by the source of the message frame (Druckman 2001). Thus, conclusion-framed review notes do appear to lead to improved effort and performance by the preparer; however, the effect of the frame does not appear to hold when the preparer is responding to a delayed review.

Lastly, participants in my study appear to engage in dysfunctional behavior in the form of over-documenting procedures performed. Specifically, one-third of the participants document performance of procedures that they did not complete. The propensity for preparers to engage in such behavior appears to be affected by reviewer delay. That is, preparers responding to timely reviews tend to over-document fewer evidence items, as a proportion of total evidence items examined, than those responding to a delayed review. This implies that, as suggested by Fox and Spector (1999), when
preparers experience anger and frustration they are more likely to behave in a way which may harm the organization.

5.3 Limitations

As with all research, my study’s limitations should be considered while interpreting the findings. One limitation is that preparers did not actually prepare the workpapers themselves. While this is held constant across conditions (and any potential impact should equally affect all conditions), it may have reduced the extent to which participants felt “ownership” of the workpapers and in turn may have influenced how they proceeded through the review note closing process. An alternative to this research design choice would be to implement the study in two rounds, one in which preparers actually perform the initial testing and prepare the workpapers, and a second round in which they receive their workpapers back with review notes. I chose not to design the study in this way for two reasons. First, I would not be able to control for the way in which the preparers performed the initial round of testing. Thus, the review of the workpapers would be different for each preparer and I would lose the experimental advantage of being able to hold all contextual features of the experiment constant other than the variables I chose to manipulate. Related to this point, this method would also require me to obtain experienced audit reviewers to perform reviews of many different participant-prepared sets of workpapers. Second, audit participants are a scarce resource and asking them to engage in two rounds of participation would have made attracting and retaining such participants more difficult.

A second potential limitation is the method in which I administered the study. Since the study calls for scarce auditor participants, the majority of participants
completed the study on their own time. Although the instructions stress the importance of working through to completion once the simulation is started, some participants occasionally appeared to have paused during the simulation. However, the computerized instrument helped identify such pauses and allowed for these instances to be taken into account while analyzing the data. Also, I was able to determine that all pauses during the completion of the task were relatively short (i.e., measured in minutes rather than hours or days). Furthermore, such work interruptions are a natural feature of the audit environment and, thus, I do not expect them to have much influence on the generalizability of the results of my study.

A third limitation is the operationalization of review timeliness. Participants did not actually wait the two/fourteen days that they are told it has taken the reviewer to return the workpapers. This was a design choice used primarily to avoid participant withdrawal from the study. In order to ensure a successful manipulation of the affect associated with time delays, participants experienced the delay by viewing a timer that counted through the passage of days, with six second intervals representing one day. In practice, preparers would most likely work on other areas of the audit and/or other audit engagements while waiting for workpapers to be reviewed. Therefore, my results relating to review timeliness must be considered in light of the extent to which frustration or anger caused by the delay might be different in this setting than in an actual audit setting. For example, the way in which preparers respond to time delays during the review process might be affected by how busy the auditor is working on other areas of the audit. If the preparer has been kept busy with other work, he or she may not experience the delay as negatively as when simply waiting for the workpapers.
Alternatively, if the preparer has moved onto another engagement and now has to rearrange his or her schedule to complete the work in the workpapers, he or she might experience the delay more negatively than if he or she is simply waiting for the workpapers.

Lastly, my study uses an experimental research design which enhances the internal validity of the study, but may reduce the external validity, or generalizability. Although I developed the simulation with the assistance of current/former auditors and audit researchers and tried to include as many realistic features as possible, the actual audit environment is naturally more complex than anything that can be included in a simulation. That is, there are other features of the audit environment that were not present during my simulation (e.g., time pressure, accountability) that might affect auditor performance. However, I have no reason to suspect that the introduction of any of these (excluded) features would differentially affect the conditions in my study.

5.4 Implications

Previous research recognizes the importance of the review process as a quality control mechanism, but has not examined the stage of the review process in which the preparer responds to and “closes” review notes. The current study will hopefully lead to an interest in further exploration of this phase of the audit to shed light upon factors that affect whether errors or issues identified during the review process are adequately resolved. Results of this study indicate that reviewer identification of an issue or error in the workpapers does not necessarily lead to detection/correction by the preparer. While it is possible that further rounds of review may uncover the lack of an adequate response by the preparer, the way in which the preparer documents his or her response to the reviewer
would be integral to such further discovery. That is, if a preparer documents no procedures related to the item in question, the reviewer is likely to detect the lack of procedures in a subsequent round of review. However, if the preparer documents that the evidence associated with an item adequately supports client treatment of that item (as happened at times in my simulation), the reviewer is not likely to detect this. Other factors, such as time pressure, other forms of affect, and other reviewer and client characteristics may be influential and should be examined to determine their impact preparer effort and performance while closing review notes.

The factors I chose to examine are (at least to some extent) under the control of the reviewer and, therefore, I expect this research to be particularly useful to audit practitioners. That is, the results of this study can help underscore the importance to audit firms and audit managers of providing timely reviews of audit workpapers and being cognizant of implications of their phrasing of review notes. With the recent spotlight on audit documentation resulting from the PCAOB’s issuance of AS3, reviewers might be more likely to provide a rationale in their review notes that highlights the importance of this audit objective. However, as Knechel (2007) warns, this focus on the formalities of the audit process, while important, may cause the preparer to view the purpose of the additional audit work as demonstrating completion of the process rather than arriving at the correct conclusions.

Also, my study contributes to the literature on dysfunctional behaviors during the audit. Previous research on this topic has examined only of self-reported measures of dysfunctional behaviors. In this study, preparers have the ability to actually engage in dysfunctional behavior (in the form of over-documentation of audit procedures) and the
software allows detection of this practice. Consistent with the results of studies using surveys and anecdotal evidence, the findings suggest that such behavior does occur in practice and identify one contextual feature of the audit (i.e., reviewer delay) that may influence the use of such behavior. Future research can examine additional factors (such as firm culture, socialization within the firm, time pressure, and other aspects of the audit that may lead to negative affect) that may contribute to this behavior and ways in which to reduce instances of over-documentation.

In addition to contributing to practice and the audit review process literature, results of this study also contribute to both the time delay and emphasis framing literatures. It adds to the time delay literature by examining the effect of delays on experienced participants in a hierarchical, time sensitive, organizational culture. For example, using students as participants, Sheldon, Thomas-Hunt, and Proell (2006) find that the negative effects of time delay are generally mitigated when the delayer is of a higher status. However, they suggest that these results may not hold for skilled participants placed in an organizational context and call for research to examine such settings. My study contributes to this literature by exploring the effect of delay in hierarchical organizational contexts and finds that time delays can have detrimental effects in such a context even when the delayer is of higher status than the delayee. This finding is contrary to results reported by Sheldon, Thomas-Hunt, and Proell (2006); however, they speculate that the status of the delayer may not mediate the impact of the delay in a time sensitive organizational setting. This research provides evidence supporting their speculation. In addition, while prior research in this area has focused primarily on self-reported attitudes and opinions (and, to a lesser extent, basic decision making), my study extends the time
delay literature by examining how participants’ reaction to a time delay affects their effort and performance levels on a relatively complex task. This research makes a similar contribution to the emphasis framing literature, which has also chiefly focused on self-reported attitudes and opinions, by examining whether an emphasis frame can affect effort and performance in a relatively complex, organizational task. This study extends the emphasis framing literature by showing that framing effects can influence task performance and by supporting the notion that the trustworthiness, or credibility, of the message framer affects the impact of the message frame.
LIST OF REFERENCES


Figure 1

Expected Interaction of Review Timeliness and Review Note Frame on Preparers’ Effort and Performance
Figure 2
*Instrument Flow*

**Task Instructions**

↓

**Client Background Information**

↓

**Submit Workpapers to Reviewer**

↓

**Experience Delay**

↓

**Receive Reviewed Workpapers**

↓

**Access Evidence, Correspond with Client, Close Review Notes**

↓

**Quit Confirmation**

↓

**Answer Post-Experimental Questions**
Figure 3

SPSS Q-Plot of Time

Normal Q-Q Plot of TotalLessTwenty

Observed Value is equal to the time from when participant received the review workpapers to when participant exits simulation and begins post-experimental questions. Pauses over 20 minutes have been extracted from the data.
APPENDIX: EXPERIMENTAL INSTRUMENT

EXHIBIT 1. Summary of Review Notes
EXHIBIT 2. Summary of Review Timeliness Manipulation
EXHIBIT 3. Description of Seeded Errors
EXHIBIT 4. Instructions and Background Information
EXHIBIT 5. Sample Workpapers Submitted/”Review in Progress” Wait Screens (i.e., Delay Counter Screens)
EXHIBIT 6. Workpapers Returned, Timely Condition
EXHIBIT 7. Workpapers Returned, Delayed Condition
EXHIBIT 8. Main Menu
EXHIBIT 9. Sample Workpaper Screens with Documentation-Framed Review Notes
EXHIBIT 10. Sample Workpaper Screens with Conclusion-Framed Review Notes
EXHIBIT 11. Firm Guidance Menu
EXHIBIT 12. Firm Guidance Menu: Sample Selection Methods
EXHIBIT 13. Firm Guidance Menu: Errors Found During Tests of Details
EXHIBIT 15. Previously Obtained Support Menu
EXHIBIT 16. Previously Obtained Support: Write-Off Memos
EXHIBIT 17. Previously Obtained Support: Write-Off Memo Example

35 Note that this appendix contains selected “PrintScreens” or “Screen Captures” of the experimental instrument. Including a sample of the possible screens accessed by participants is necessary because of the nature of the instrument. Once participants navigate past the screen in which the workpapers are returned from the reviewer, the subsequent screen they encounter is dependent upon their navigation choices. Note that all screens containing manipulations and all post-experimental questions are included in the subset presented.

EXHIBIT 19. Previously Obtained Support: Late Cut-Off Detail

EXHIBIT 20. Previously Obtained Support: Early Cut-Off Detail

EXHIBIT 21. Previously Obtained Support: AR Detail

EXHIBIT 22. Communication with Client Personnel

EXHIBIT 23. Sample Communication with Client Personnel, Timely Condition

EXHIBIT 24. Sample Communication with Client Personnel, Delayed Condition

EXHIBIT 25. Invoice Example

EXHIBIT 26. Shipping Document Example

EXHIBIT 27. Customer Correspondence Example

EXHIBIT 28. Quit Confirmation Screen and Post-Experimental Questions
**EXHIBIT 1. Summary of Review Notes**  
*(with frame manipulation in bold, italics where applicable)*

<table>
<thead>
<tr>
<th>Review Note #</th>
<th>Documentation Frame</th>
<th>Conclusion Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Please adjust date in the procedures to include 12/31/08.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Number does not agree to number on “Summary GL” tab. Number here should agree to total of the detail from which you made selections.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Please perform more procedures on the unconfirmed selections (denoted by your tickmark b). See the audit program for guidance (step 1F). We need to make sure all of the necessary procedures are documented.</td>
<td>Please perform more procedures on the unconfirmed selections (denoted by your tickmark b). See the audit program for guidance (step 1F). We need to make sure that all of these receivables are valid.</td>
</tr>
<tr>
<td>4</td>
<td>Please check firm guidance to make sure that a haphazard sample is appropriate for 10 selections.</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Please check to see that write-offs were properly authorized. We need to make sure that all of the steps on the audit program have been documented.</td>
<td>Please check to see that write-offs were properly authorized. We need to make sure that we arrive at the proper conclusion regarding their authorization assertion.</td>
</tr>
<tr>
<td>6</td>
<td>We still need to perform additional cut-off procedures due to this error (see firm guidance on errors found during test of details). We need to make sure our documentation holds up.</td>
<td>We still need to perform additional cut-off procedures due to this error (see firm guidance on errors found during test of details). We need to make sure our conclusions (regarding cut-off) hold up.</td>
</tr>
<tr>
<td>7</td>
<td>We use PY balance to set our expectation of the allowance. Make sure that this is an appropriate basis for our expectation or see if you can come up with a better one. We really need to make sure that we document a justifiable expectation.</td>
<td>We use PY balance to set our expectation of the allowance. Make sure that this is an appropriate basis for our expectation or see if you can come up with a better one. We really need to make sure we use an expectation that adequately tests the appropriateness of the allowance.</td>
</tr>
</tbody>
</table>
**EXHIBIT 2. Summary of Review Timeliness Manipulation**  
(with differences in bold, italics where applicable)

<table>
<thead>
<tr>
<th>Location</th>
<th>Timely</th>
<th>Delayed</th>
</tr>
</thead>
<tbody>
<tr>
<td>After participant submits workpapers to the reviewer. “It has been X day(s) since you sent the workpapers to Bill for review.”</td>
<td>“X” counts up from 0 to 2 in intervals of 6 seconds.</td>
<td>“X” counts up from 0 to <strong>14</strong> in intervals of 6 seconds.</td>
</tr>
<tr>
<td>When the workpapers are returned</td>
<td>“You have received the reviewed workpapers back from Bill after he has had them for 2 days.”</td>
<td>“You have received the reviewed workpapers back from Bill after he has had them for <strong>14</strong> days.”</td>
</tr>
<tr>
<td>The first time participant accesses links to supporting invoices for subsequent receipts testing through Meg Black</td>
<td>“You have provided Ms Black the invoice numbers of the unconfirmed accounts receivable. She said she thought you had finished testing receivables a couple of <strong>days</strong> ago but that it was no problem to pull the invoices for you.”</td>
<td>“You have provided Ms Black the invoice numbers of the unconfirmed accounts receivable. She said she thought you had finished testing receivables a couple of <strong>weeks</strong> ago but that it was no problem to pull the invoices for you.”</td>
</tr>
<tr>
<td>The first time participant accesses links to shipping documents for subsequent receipts testing through John Sawyer</td>
<td>“You have provided the invoice numbers of the unconfirmed accounts receivable to Mr Sawyer. He mentioned that he thought you had finished testing AR a few <strong>days</strong> ago. He has provided you the shipping documents for the accounts which he manages.”</td>
<td>“You have provided the invoice numbers of the unconfirmed accounts receivable to Mr Sawyer. He mentioned that he thought you had finished testing AR a few <strong>weeks</strong> ago. He has provided you the shipping documents for the accounts which he manages.”</td>
</tr>
<tr>
<td>The first time participant accesses links to shipping documents for subsequent receipts testing through Kate Burke</td>
<td>“You have provided to Ms Burke the invoice numbers of the unconfirmed accounts receivable. She expressed surprise that you were still testing Accounts Receivable because she thought you had finished with that a couple of <strong>days</strong> ago. She has provided you with the shipping documents for the accounts</td>
<td>“You have provided to Ms Burke the invoice numbers of the unconfirmed accounts receivable. She expressed surprise that you were still testing Accounts Receivable because she thought you had finished with that a couple of <strong>weeks</strong> ago. She has provided you with the shipping documents for the accounts</td>
</tr>
<tr>
<td>The first time participant investigates previously noted cut-off error with Kate Burke.</td>
<td>“Kate Burke reiterated and expanded upon the exploration for the error that you found in the cut-off testing a couple of days ago.”</td>
<td>“Kate Burke reiterated and expanded upon the exploration for the error that you found in the cut-off testing a couple of weeks ago.”</td>
</tr>
<tr>
<td>The first time participant investigates previously noted cut-off error with Meg Black.</td>
<td>“Meg Black reiterated and expanded upon the exploration for the error that you found in the cut-off testing a couple of days ago.”</td>
<td>“Meg Black reiterated and expanded upon the exploration for the error that you found in the cut-off testing a couple of weeks ago.”</td>
</tr>
</tbody>
</table>
**EXHIBIT 3. Description of Seeded Errors**

<table>
<thead>
<tr>
<th>Corresponding Review Note #</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Invoice #41756 reads $112,735.20. The selection is listed in the workpapers and in the client’s account balance detail as $112,375.20.</td>
</tr>
<tr>
<td>3</td>
<td>Invoice #41965 reads $278,483.13. The selection is listed in the workpapers and in the client’s account balance detail as $287,483.13.</td>
</tr>
<tr>
<td>3</td>
<td>The shipping document provided to correspond with invoice number #41786 is for purchase memo #34096. Invoice #41786 is for purchase memo #34131. Also, the date on the shipping document does not match the date on the invoice and the merchandise listed on the two documents does not match. If the participant asks the client for the correct shipping document he/she is told it is missing, that a copy will be requested from the shipping company, but it isn’t expected for a couple of weeks.</td>
</tr>
<tr>
<td>3</td>
<td>If the participant asks the account manager about customer correspondence for “Olivier’s Olive Oil,” he/she is told the client will contact the customer and to come back later. If the participant returns before waiting 2 minutes, he/she is told that the client is still waiting to hear from the customer. If the participant returns after waiting 5 minutes, he/she is told that the merchandise for that invoice had been returned prior to year end and that the amount is not a valid receivable.</td>
</tr>
<tr>
<td>3</td>
<td>If the participant asks the account manager about customer correspondence for “alligator aid,” he/she is told the merchandise listed on invoice #41900 was returned and should have been offset by credit memo #3045.</td>
</tr>
<tr>
<td>5</td>
<td>Credit memo #2178 is missing the account manager’s signature</td>
</tr>
<tr>
<td>5</td>
<td>Credit memo #2141 contains the wrong account manager’s signature (Kate Burke signed instead of John Sawyer)</td>
</tr>
<tr>
<td>6</td>
<td>Credit memo #3048 wasn’t issued or recorded until after 12/31/08, but it relates to merchandise in which ownership transferred back to the client before 12/31/08.</td>
</tr>
</tbody>
</table>
EXHIBIT 4. Instructions and Background Information

Instructions

Thank you for participating in this research study. It should take approximately 1 hour to complete this simulated audit case. It is important to complete this exercise in one sitting without interruption. If you do not feel as though you currently have the time to complete this in one sitting, please click the "come back" button below and return to this task at a later time.

Come Back Later

The task will involve opening PDF files and Excel documents. So to avoid confusion, please close all other documents and applications on your computer before proceeding.

This study asks you to assume the role of a staff auditor and involves submitting workpapers and later responding to review notes. When you submit the workpapers, you may experience a short delay while your reviewer reviews the workpaper. The workpapers are in the form of Microsoft Excel spreadsheets. Any review notes will be placed in the workpapers using the "comments" function. Please address the review notes exactly as you would if you had received these review notes from your reviewer at an actual client of yours. You can make changes to the workpapers throughout the Excel spreadsheet and leave comments to the reviewer in the review notes themselves. Unless specifically asked in a review note to perform additional testing, please simply document any additional issues or errors you find, if you find any.

You may want to have a sheet of paper and pencil handy as you proceed in order to take notes as you would in practice.

Please press the "continue" button below to proceed to review the client's background information.

Continue
Background

Assume that you are an auditor assigned to the 12/31/02 fiscal year-end audit of Sprandell Inc., a publicly held, mid-sized manufacturer of plastic bottles. Sprandell Inc. sells their plastic bottles to a variety of customers for use in the packaging of a wide assortment of products such as food and drink, automotive oil, and personal care products.

One of the areas you have been assigned to test is Accounts Receivable. You have three main client contacts for testing Accounts Receivable at Sprandell Inc.: (1) Meg Black, the Accounts Receivable Supervisor, (2) Kate Burke, Account Manager, and (3) John Sawyer, Account Manager. Meg Black oversees the accounting for Accounts Receivable, and the Account Managers are responsible for providing customer service and correspondence with specific accounts. The signatures of both the Accounts Receivable Supervisor (Meg Black) and the Account Manager responsible for the specific account (either Kate Burke or John Sawyer, depending on the customer) are necessary for the write-off of an account receivable.

The internal audit function of Sprandell Inc. is considered independent of the accounting function. You can treat any evidence obtained from or verified by internal audit as adequate audit evidence.

Assume you have just completed the testing of Accounts Receivable and are ready to send the workpapers to Manager Bill Davis for review. Once Bill has completed his review, you will be able to access the reviewed workpapers. You expect Bill to take approximately 2 days to review the workpapers and return them to you.

Press the ‘submit workpapers’ button to send the workpapers to Bill.

Submit Workpapers
EXHIBIT 5. Sample Workpapers Submitted/"Review in Progress" Wait Screens (i.e., Delay Counter Screens)

Workpapers Submitted: Review In Process

It has been 0 day(s) since you sent the workpapers to Bill for review.

You will be notified when the workpapers are returned.

Expected review time: 2 days
Workpapers Submitted: Review In Process

It has been 1 day(s) since you sent the workpapers to Bill for review.

You will be notified when the workpapers are returned.

Expected review time: 2 days
EXHIBIT 6. Workpapers Returned, Timely Condition

Workpapers Returned

You have received the reviewed workpapers back from Bill after he has had them for 2 days.

Expected review time: 2 days

To access the reviewed workpapers, please click on the "Main Menu" button below.

From the Main Menu you can also access the following: (1) evidence that you had previously gathered for Accounts Receivable testing, (2) excerpts from firm guidance, and (3) key client personnel from whom you can request additional audit evidence. You can also return to the instructions and background information.

Main Menu
EXHIBIT 7. Workpapers Returned, Delayed Condition

Workpapers Returned

You have received the reviewed workpapers back from Bill after he has had them for 14 days.

Expected review time: 2 days

To access the reviewed workpapers, please click on the "Main Menu" button below.

From the Main Menu you can also access the following: (1) evidence that you had previously gathered for Accounts Receivable testing, (2) excerpts from firm guidance, and (3) key client personnel from whom you can request additional audit evidence. You can also return to the instructions and background information.
Exhibit 8. Main Menu

Main Menu

- **Workpapers**
  To examine the reviewed workpapers and review notes

- **Guidance**
  To view excerpts of relevant firm guidance and the Accounts Receivable audit program

- **Support**
  To view supporting documents that you obtained previously for testing (i.e., write-off testing memos, account detail, etc.)

- **Client Personnel**
  To communicate with relevant Client Personnel to obtain additional information and evidence

- **Instructions**
  To return to the Instructions / Background Information Screens

- **Finished**
  Review Notes closed – proceed to brief post-task questionnaire
EXHIBIT 9. Sample Workpaper Screens with Documentation-Framed Review Notes

<table>
<thead>
<tr>
<th>Customer</th>
<th>Invoice #</th>
<th>Date</th>
<th>$ Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sinus Aide</td>
<td>41569</td>
<td>9/27/2008</td>
<td>45,767.90</td>
</tr>
<tr>
<td>Jay &amp; Jack's Auto Oil</td>
<td>41624</td>
<td>10/9/2008</td>
<td>8,152,958.95</td>
</tr>
<tr>
<td>Generic Cola Inc.</td>
<td>41687</td>
<td>10/19/2008</td>
<td>8,264.65</td>
</tr>
<tr>
<td>Florda Juice Corp.</td>
<td>41699</td>
<td>11/3/2008</td>
<td>68,585.40</td>
</tr>
<tr>
<td>Mars Liquid Detergents</td>
<td>41790</td>
<td>11/6/2008</td>
<td>745,714.15</td>
</tr>
<tr>
<td>Silver Springs</td>
<td>41711</td>
<td>11/9/2008</td>
<td>45,825.31</td>
</tr>
<tr>
<td>Petro Oil</td>
<td>41768</td>
<td>11/14/2008</td>
<td>112,376.20</td>
</tr>
<tr>
<td>Jay &amp; Jack's Auto Oil</td>
<td>41764</td>
<td>11/17/2008</td>
<td>6,144,753.61</td>
</tr>
<tr>
<td>Tex Oil</td>
<td>41767</td>
<td>11/20/2008</td>
<td>1,782,960.59</td>
</tr>
<tr>
<td>Penn Oil</td>
<td>41768</td>
<td>11/23/2008</td>
<td>57,459.00</td>
</tr>
<tr>
<td>The Bath &amp; Body Shop</td>
<td>41765</td>
<td>11/30/2008</td>
<td>412,426.15</td>
</tr>
<tr>
<td>Fresh Squeezed</td>
<td>41788</td>
<td>12/1/2008</td>
<td>25,470.90</td>
</tr>
<tr>
<td>Olin's Olive Oil</td>
<td>41811</td>
<td>12/4/2008</td>
<td>12,335.37</td>
</tr>
<tr>
<td>Tex Oil</td>
<td>41845</td>
<td>12/9/2008</td>
<td>3,785,490.00</td>
</tr>
<tr>
<td>Tex Oil</td>
<td>41877</td>
<td>12/9/2008</td>
<td>780,065.12</td>
</tr>
<tr>
<td>Windowx</td>
<td>41937</td>
<td>12/9/2008</td>
<td>311,411.42</td>
</tr>
<tr>
<td>Soup to Go</td>
<td>41888</td>
<td>12/11/2008</td>
<td>56,287.11</td>
</tr>
<tr>
<td>Liquid Chocolate</td>
<td>41889</td>
<td>12/14/2008</td>
<td>89,543.50</td>
</tr>
<tr>
<td>Drive R Us</td>
<td>41936</td>
<td>12/15/2008</td>
<td>98,945.25</td>
</tr>
<tr>
<td>Alligator Air</td>
<td>41990</td>
<td>12/19/2008</td>
<td>261.71</td>
</tr>
<tr>
<td>Pure Oil</td>
<td>41993</td>
<td>12/20/2008</td>
<td>389,545.15</td>
</tr>
<tr>
<td>Diet Water</td>
<td>41965</td>
<td>12/27/2008</td>
<td>287,483.13</td>
</tr>
<tr>
<td>Ann's Aromatherapy</td>
<td>41990</td>
<td>12/29/2008</td>
<td>933,652.00</td>
</tr>
<tr>
<td>Soup to Go</td>
<td>42091</td>
<td>12/29/2008</td>
<td>247,684.30</td>
</tr>
<tr>
<td>Vat O Wine</td>
<td>42092</td>
<td>12/29/2008</td>
<td>3,456,784.42</td>
</tr>
</tbody>
</table>
## Total Write-offs

*From "Rollforward" Tab*

<table>
<thead>
<tr>
<th>Customer</th>
<th>Memo #</th>
<th>$ Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jag O Wine</td>
<td>2110</td>
<td>170,130.00</td>
</tr>
<tr>
<td>Cold &amp; Sinus Care</td>
<td>2125</td>
<td>88,141.09</td>
</tr>
<tr>
<td>Sunset Soap</td>
<td>2130</td>
<td>11,575.12</td>
</tr>
<tr>
<td>Petrol UK</td>
<td>2141</td>
<td>6,570.60</td>
</tr>
<tr>
<td>Sunset Soap</td>
<td>2145</td>
<td>5,330.33</td>
</tr>
<tr>
<td>Sunset Soap</td>
<td>2150</td>
<td>256.72</td>
</tr>
<tr>
<td>Jay &amp; Jack’s Auto Oil</td>
<td>2178</td>
<td>9,555.37</td>
</tr>
<tr>
<td>Scent of a Pet</td>
<td>2181</td>
<td>7,180.10</td>
</tr>
<tr>
<td>Scent of a Pet</td>
<td>2182</td>
<td>53,014.46</td>
</tr>
<tr>
<td>Scent of a Pet</td>
<td>2183</td>
<td>8,986.56</td>
</tr>
</tbody>
</table>

---

**Tickmarks**

1. Number of necessary selections was calculated based on materiality and risk level. See manual workpapers for calculation.
2. Traced selections to supporting documentation and correspondence with the customer that indicated amount is uncollectible.
Based on review of supporting documentation and discussion with management, credit amounts are due to customer overpayments. Traced and agreed amounts to supporting documentation. The customer has been informed of the overpayment and will apply the credit to future purchases of the customer. Amount has been reclassified as a liability. Please see "Other Liabilities" testing for further information.

Amounts were traced and agreed to shipping/receiving records and credit memo. Amount appears correctly classified as a credit for 2008.

Amounts were traced and agreed to shipping/receiving records and credit memo. Amount appears correctly classified as a credit for 2009.

Amount was traced and agreed to shipping/receiving records and credit memo. Amount relates to returned merchandise in which ownership was transferred back to Spoanal Inc. on the last business day of 2008. Due to the clear immateriality of the amount, no further investigation is warranted.

Review Note 6 of 7: We still need to perform additional cut-off procedures due to this error (see firm guidance on errors found during test of details). We need to make sure our documentation holds up.
**Purpose:** The purpose of this workpaper is to assess the reasonableness of the allowance for doubtful accounts (AR Reserve) balance as of 12/31/08. Please see additional testing in the manual and interim workpapers.

**Procedures:** Analytical procedures were used to assess the adequacy of the allowance (reserve) account as of 12/31/08. Fluctuations greater than 5% of the recorded account balance will be investigated.

**Conclusion:** Based on work performed during interim testing and at this workpaper. Accounts Receivable appear to be reasonably stated.

<table>
<thead>
<tr>
<th>12/31/08 Actual Balance</th>
<th>12/31/08 Expected Balance</th>
<th>Fluctuation</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allowance</td>
<td>5,376,644</td>
<td>5,640,758</td>
<td>(264,114)</td>
</tr>
</tbody>
</table>

**Tickmarks**

- **m:** Prior year balance was used as the expectation for the current year.
- **n:** Amount obtained from prior year workpapers.
- **o:** Fluctuation is within established threshold.
- **p:** Based on firm guidance, the threshold is calculated as 5% of the expected balance.

**Review Note 7 of P:** We use FY Balance to set our expectation of the allowance. Make sure that this is an appropriate basis for our expectation or see if you can come up with a better one. We really need to make sure that we document a purposefully expectation.
EXHIBIT 10. Sample Workpaper Screens with Conclusion-Framed Review Notes
Total Write-offs (3,690.427) Fm "Rollforward" Tab

Total Number of Selections 10

g

<table>
<thead>
<tr>
<th>Customer</th>
<th>Memo #</th>
<th>$ Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jug O Wine</td>
<td>2118</td>
<td>170,130.09</td>
</tr>
<tr>
<td>Cold &amp; Sinus Care</td>
<td>2125</td>
<td>60,141.09</td>
</tr>
<tr>
<td>Sunset Soap</td>
<td>2139</td>
<td>11,575.15</td>
</tr>
<tr>
<td>Petrol UK</td>
<td>2141</td>
<td>6,670.60</td>
</tr>
<tr>
<td>Sunset Soap</td>
<td>2149</td>
<td>5,238.33</td>
</tr>
<tr>
<td>Sunset Soap</td>
<td>2155</td>
<td>2,082.72</td>
</tr>
<tr>
<td>Jay &amp; Jack's Auto Oil</td>
<td>2178</td>
<td>9,626.57</td>
</tr>
<tr>
<td>Scent of a Pet</td>
<td>2181</td>
<td>76,106.10</td>
</tr>
<tr>
<td>Scent of a Pet</td>
<td>2182</td>
<td>53,521.40</td>
</tr>
<tr>
<td>Scent of a Pet</td>
<td>2183</td>
<td>8,668.56</td>
</tr>
</tbody>
</table>

Review Note 5 of 7: Please check to see that write-offs were properly authorized. We need to make sure that we arrive at the proper conclusion regarding their authorization assertion.

Tickmarks

g

Number of necessary selections was calculated based on materiality and risk level. See manual workpapers for calculation.

h

Traced selections to supporting documentation and correspondence with the customer that indicated amount is uncollectible.
Based on review of supporting documentation and discussion with management, credit amounts are due to customer overpayments. Traced and agreed amounts to supporting documentation. The customer has been informed of the overpayment and will apply the credit to future purchases of the customer. Amount has been reclassified as a liability. Please see "Other Liabilities" testing for further information.

Amounts were traced and agreed to shipping/receiving documents and credit memo. Amount appears correctly classified as a credit for 2008.

Amounts were traced and agreed to shipping/receiving documents and credit memo. Amount appears correctly classified as a credit for 2009.

Amount was traced and agreed to shipping/receiving records and credit memo. Amount relates to returned merchandise in which ownership was transferred back to Spirandol Inc. on the last business day of 2008. Due to the clear immateriality of the amount, no further investigation is warranted.

Review Note 6 of 7: we still need to perform additional cut-off procedures due to this error (see firm guidance on errors found during test of details). We need to make sure our conclusions (regarding cut-off) hold up.
**PURPOSE:** The purpose of this workpaper is to assess the reasonableness of the allowance for doubtful accounts (4R Reserve) balance as of 12/31/08. Please see additional testing in the manual and interim workpapers.

**PROCEDURES:** Analytical procedures were used to assess the adequacy of the allowance (reserve) account as of 12/31/08. Fluctuations greater than 5% of the recorded account balance will be investigated.

**CONCLUSION:** Based on work performed during interim testing and at this workpaper, Accounts Receivable appear to be reasonably stated.

<table>
<thead>
<tr>
<th>12/31/08 Actual Balance</th>
<th>12/31/08 Expected Balance</th>
<th>Fluctuation</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allowance</td>
<td>5,376,644</td>
<td>m (254,114)</td>
<td>268,832</td>
</tr>
</tbody>
</table>

**Tickmarks**

- **m** Prior year balance was used as the expectation for the current year. Amount obtained from prior year workpapers.
- **n** Fluctuation is within established threshold.
- **o** Based on firm guidance, the threshold is calculated as 5% of the expected balance.

**Review Note 7 of F:**

We use PV Balance to set our expectation of the allowance. Make sure that there's an appropriate basis for our expectation or see if you can come up with a better one. We really need to make sure we use an expectation that adequately tests the appropriateness of the allowance.
EXHIBIT 11. Firm Guidance Menu

Guidance Menu

Sample Selection  
To view firm guidance related to sample selection methods

Test of Details - Errors  
To view firm guidance related to errors found during detailed testing

AF Audit Program  
To view the Accounts Receivable audit program

Back to Main Menu
EXHIBIT 12. Firm Guidance Menu: Sample Selection Methods

Sample Selection Methods

When possible, audit software should be used to obtain a statistical sample when performing tests of details or other similar tests. If such software cannot be used, a statistical sample should be obtained using some other available method. When the required number of selections is 10 or less, a non-statistical or haphazard sample may be used at the discretion of members of the engagement team.

Back to Firm Guidance Menu

Back to Main Menu
EXHIBIT 13. Firm Guidance Menu: Errors Found During Tests of Details

Errors Found During Tests of Details

When an error is found during detailed testing, the auditor should make additional selections for testing (regardless of the dollar amount of the error). The auditor should attempt to obtain an understanding of how the error occurred and determine whether it appears to be an isolated instance, or whether it is indicative of a more pervasive problem. If the error can be isolated to a particular type of transaction, additional testing should be focused on other transactions of this type. For example, if the error was due to specifics related to a particular customer or vendor, additional testing should focus on transactions involving that customer or vendor.

If the error cannot be isolated to a particular type of transaction, additional testing should cover the entire account balance.

Back to Firm Guidance Menu

Back to Main Menu

EXHIBIT 15. Previously Obtained Support Menu
Previously Obtained Support

To access the credit memos you previously obtained for write-off testing at the "Write-Offs" tab of the workpapers.

To access the detailed lists from which you chose your selections for cut-off testing. The lists were obtained from the client.

To access the detailed list from which you chose your selections for Accounts Receivable existence testing. The list was obtained from the client.
EXHIBIT 16. Previously Obtained Support: Write-Off Memos

Please click on the write-off memo number below to access the corresponding write-off memo that you previously obtained from the Account Managers.

2110  2155
2125  2178
2139  2181
2141  2182
2149  2183

Back to Previous Support Menu

Back to Main Menu
EXHIBIT 17. Previously Obtained Support: Write-Off Memo Example

![Write-Off Memo Example](Exhibit17.png)

Cut-Off Testing Detail

Late Cut-Off Detail
To access the detailed list of all credit memo items issued by Sprandel Inc for the 15 days after fiscal year end

Early Cut-Off Detail
To access the detailed list of all credit memo items issued by Sprandel Inc for the 15 days prior to fiscal year end

Back to Previous Support

Back to Main Menu
<table>
<thead>
<tr>
<th>CUSTOMER</th>
<th>AMOUNT</th>
<th>MEMO</th>
<th>DATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mars Liquid Detergents</td>
<td>5,879.00</td>
<td>3046</td>
<td>01/02/2009</td>
</tr>
<tr>
<td>Silver Springs</td>
<td>114,741.03</td>
<td>3047</td>
<td>01/03/2009</td>
</tr>
<tr>
<td>Mars Liquid Detergents</td>
<td>5,754.46</td>
<td>3048</td>
<td>01/04/2009</td>
</tr>
<tr>
<td>Generic Juice</td>
<td>4,956.88</td>
<td>3049</td>
<td>01/08/2009</td>
</tr>
<tr>
<td>Alligator Aid</td>
<td>10,158.50</td>
<td>3050</td>
<td>01/11/2009</td>
</tr>
<tr>
<td>Drinks R Us</td>
<td>369,765.44</td>
<td>3051</td>
<td>01/12/2009</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>511,255.31</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### EXHIBIT 20. Previously Obtained Support: Early Cut-Off Detail

<table>
<thead>
<tr>
<th>CUSTOMER</th>
<th>AMOUNT</th>
<th>MEMO</th>
<th>DATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penns Oil</td>
<td>96,541.67</td>
<td>3041</td>
<td>12/17/2008</td>
</tr>
<tr>
<td>Soupy Soap</td>
<td>4,141.01</td>
<td>3042</td>
<td>12/20/2008</td>
</tr>
<tr>
<td>Soupy Soap</td>
<td>313.55</td>
<td>3043</td>
<td>12/22/2008</td>
</tr>
<tr>
<td>The Bath &amp; Body Shop</td>
<td>325,845.12</td>
<td>3044</td>
<td>12/26/2008</td>
</tr>
<tr>
<td>Alligator Aid</td>
<td>75,987.34</td>
<td>3045</td>
<td>12/30/2008</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>402,828.69</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### EXHIBIT 21. Previously Obtained Support: AR Detail

Note: This list was obtained from the client and represents the gross AR detail. Schedule was footed without exception. See testing of selected items at workpaper tab "AR Existence."

<table>
<thead>
<tr>
<th>Invoice #</th>
<th>Customer</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>41569</td>
<td>Sinus Aide</td>
<td>45,767.90</td>
</tr>
<tr>
<td>41590</td>
<td>Vat O Grease</td>
<td>83,473.90</td>
</tr>
<tr>
<td>41593</td>
<td>Petrol TX</td>
<td>109,053.12</td>
</tr>
<tr>
<td>41594</td>
<td>Fresh Squeezed</td>
<td>2,178.90</td>
</tr>
<tr>
<td>41599</td>
<td>Anna’s Aromatherapy</td>
<td>455.67</td>
</tr>
<tr>
<td>41600</td>
<td>Winbox</td>
<td>567,145.90</td>
</tr>
<tr>
<td>41604</td>
<td>Liquid Chocolate</td>
<td>78,654.00</td>
</tr>
<tr>
<td>41608</td>
<td>Petrol TX</td>
<td>318,976.10</td>
</tr>
<tr>
<td>41610</td>
<td>Pure Oil</td>
<td>1,950,875.31</td>
</tr>
<tr>
<td>41615</td>
<td>Vat O Grease</td>
<td>821,096.15</td>
</tr>
<tr>
<td>41616</td>
<td>Vat O Wine</td>
<td>75,623.41</td>
</tr>
<tr>
<td>41621</td>
<td>Vegemator</td>
<td>10,351.67</td>
</tr>
<tr>
<td>41624</td>
<td>Jay &amp; Jack’s Auto Oil</td>
<td>8,152,858.95</td>
</tr>
<tr>
<td>41625</td>
<td>Neuman’s Own</td>
<td>87,563.25</td>
</tr>
<tr>
<td>41630</td>
<td>Petrol TX</td>
<td>110,432.83</td>
</tr>
<tr>
<td>41634</td>
<td>Catty Catsup</td>
<td>334,765.23</td>
</tr>
<tr>
<td>41637</td>
<td>Fleur de Moutard</td>
<td>98,128.55</td>
</tr>
<tr>
<td>41638</td>
<td>Huggy Honey</td>
<td>590.43</td>
</tr>
<tr>
<td>41641</td>
<td>Sinus Aid &amp; Sinus Care</td>
<td>21,384.72</td>
</tr>
<tr>
<td>41658</td>
<td>Anna’s Aromatherapy</td>
<td>51,983.72</td>
</tr>
<tr>
<td>41662</td>
<td>Liquid Chocolate</td>
<td>296,404.82</td>
</tr>
<tr>
<td>41663</td>
<td>Squeeze &amp; Scrub</td>
<td>523,307.63</td>
</tr>
</tbody>
</table>
EXHIBIT 22. Communication with Client Personnel

Client Personnel

Meg Black, Accounts Receivable Supervisor, approves write-offs, calculates the AR Allowance, drafts and books AR-related journal entries, and keeps all invoices. The Account Managers, Kate Burke and John Sawyer, are responsible for individual accounts. They initiate and sign-off on write-offs, correspond with customers, and keep all shipping documents.

---

**Meg Black**

To communicate with Meg Black, AR Supervisor.

**Kate Burke**

To communicate with Kate Burke, Account Manager for Cold & Scurvy Care, Florida Juice Corp, Generic Cola Inc., Joy & Jack’s Auto Oil, Mars Liquid Detergents, Oliver’s Olive Oil, Sunset Soap, Soup to Go, and others.

**John Sawyer**

To communicate with John Sawyer, Account Manager for Alligator Aid, The Bath & Body Shop, Diet Water, Jug O’ Wine, Penns Oil, Petrol UK, Scents of a Pet, Silver Springs, and others.

---

**Main Menu**
Meg Black

Please use 2 to 3 words below to describe one issue you would like to discuss with Meg Black, Accounts Receivable Supervisor. For example, if you wanted to request support for testing cash accounts (as opposed to AR), you might type “cash testing”, “bank reconciliations”, or “bank statements”. You will be able to ask her about more than one item. However, please inquire about one item at a time.

Please press "enter" after typing your words.

[Blank field]

Back to Main Menu
Kate Burke

Please use 2 to 3 words below to describe one issue you would like to discuss with Kate Burke, Account Manager. For example, if you wanted to request support for testing of cash accounts (as opposed to AR), you might type "cash testing", "bank reconciliations", or "bank statements". You will be able to ask her about more than one item. However, please inquire about one item at a time.

Please press "enter" after typing your words.
John Sawyer

Please use 2 to 3 words below to describe one issue you would like to discuss with John Sawyer, Account Manager. For example, if you wanted to request support for testing cash accounts (as opposed to AR), you might type "cash testing", "bank reconciliations" or "bank statements". You will be able to ask him about more than one item. However, please inquire about one item at a time.

Please press "enter" after typing your words.

Back to Main Menu
EXHIBIT 23. Sample Communication with Client Personnel, Timely Condition*

Accounts Receivable Invoices

You have provided Ms Black the invoice numbers of the unconfirmed accounts receivable. She said she thought you had finished testing receivables a couple of days ago but that it was no problem to pull the invoices for you. She suggested that, if necessary, you should ask the specific account managers (Kate Burke and John Sawyer) for shipping documents and for information on customer correspondence.

Please click on the invoice number below to access the corresponding invoice.

41687          41811
41699          41888
41756          41900
41768          41965
41786          42001

*The reference to time noted on screens in Exhibit 23 is only present the first time these screens are accessed.
Accounts Receivable Shipping Documents

You have provided to Ms. Burke the invoice numbers of the unconfirmed accounts receivable. She expressed surprise that you were still testing Accounts Receivable because she thought you had finished with that a couple of days ago. She has provided you with the shipping documents for the accounts which she manages. She informed you that John Sawyer manages the remaining accounts and Meg Black, the Accounts Receivable supervisor, keeps the invoices until they are paid.

Please click on the invoice number below to access the corresponding shipping document.

41687
41699
41811
41888
42001

Ask Ms. Burke Another Question

Back to Main Menu
Accounts Receivable Shipping Documents

You have provided the invoice numbers of the unconfirmed accounts receivable to Mr. Sawyer. He mentioned that he thought you had finished testing AR a few days ago. He has provided you the shipping documents for the accounts which he manages. He informed you that Kate Burke manages the remaining accounts and Meg Black, the Accounts Receivable supervisor, keeps the invoices until they are paid.

Please click on the invoice number below to access the corresponding shipping document.

- 41756
- 41768
- 41786
- 41900
- 41965

Ask Mr. Sawyer Another Question

Back to Main Menu
Cut-Off Testing

Kate Burke reiterated and expanded upon the explanation for the error that you found in the cut-off testing a couple of days ago. She explained that Mars Liquid Detergents has different terms for returning merchandise than the rest of Sprandel Inc.’s customers. The title of the merchandise changes hands (for Mars Liquid Detergents) when the carrier picks up the merchandise. For all of Sprandel’s other customers, the title of the merchandise does not change hands until the merchandise is delivered by the carrier. The merchandise was picked up by the carrier on 12/30/08 but was not delivered to Sprandel Inc. until 1/2/09. By mistake, a credit memo was not issued until 1/2/09 because that is the process for other Sprandel Inc. customers. Kate has indicated that, if you think it’s necessary, Meg Black will discuss adjusting for the error with your manager Bill Davis.

Ms Burke has indicated that she will provide any further support you require for additional credit memo cut-off testing for the accounts she manages. She suggested that you request the support by credit memo number. She said she’s pretty sure Meg previously supplied you with the cut-off detail.

Ask Ms Burke Another Question

Back to Main Menu
Cut-Off Testing

Meg Black reiterated and expanded upon the explanation for the error that you found in the cut-off testing a couple of days ago. She explained that Mars Liquid Detergent has different terms for returning merchandise than the rest of Sprandel Inc.'s customers. The title of the merchandise changes hands (for Mars Liquid Detergent) when the carrier picks up the merchandise. For all of Sprandel's other customers, the title of the merchandise does not change hands until the merchandise is delivered by the carrier. The merchandise was picked up by the carrier on 12/30/08 but was not delivered to Sprandel Inc. until 1/2/09. By mistake, a credit memo was not issued until 1/2/09 because that is the process for other Sprandel Inc. customers. Meg has indicated that, if you feel it's necessary, she will discuss adjusting for the error with your manager Bill Davis.

Ms. Black also mentioned that you already have the detail for any additional cut-off testing you may want to perform and suggested any further support related to additional credit memo cut-off testing should be obtained from the Account Managers.

[Buttons: Ask Ms. Black Another Question, Back to Main Menu]
EXHIBIT 24. Sample Communication with Client Personnel, Delayed Condition*

*The reference to time noted on screens in Exhibit 24 is only present the first time these screens are accessed.

Accounts Receivable Invoices

You have provided Mr. Black the invoice numbers of the unconfirmed accounts receivable. She said she thought you had finished testing receivables a couple of weeks ago but that it was no problem to pull the invoices for you. She suggested that, if necessary, you should ask the specific account managers (Kate Burke and John Sawyer) for shipping documents and for information on customer correspondence.

Please click on the invoice number below to access the corresponding invoice:

<table>
<thead>
<tr>
<th>41687</th>
<th>41811</th>
</tr>
</thead>
<tbody>
<tr>
<td>41699</td>
<td>41888</td>
</tr>
<tr>
<td>41756</td>
<td>41900</td>
</tr>
<tr>
<td>41768</td>
<td>41965</td>
</tr>
<tr>
<td>41786</td>
<td>42001</td>
</tr>
</tbody>
</table>

Ask Mr. Black Another Question

Back to Main Menu
Accounts Receivable Shipping Documents

You have provided to Ms. Burke the invoice numbers of the unconfirmed accounts receivable. She expressed surprise that you were still testing Accounts Receivable because she thought you had finished with that a couple of weeks ago. She has provided you with the shipping documents for the accounts which she manages. She informed you that John Sawyer manages the remaining accounts and Meg Black, the Accounts Receivable supervisor, keeps the invoices until they are paid.

Please click on the invoice number below to access the corresponding shipping document.

41687
41699
41811
41888
42001

Ask Ms. Burke Another Question
Back to Main Menu
Accounts Receivable Shipping Documents

You have provided the invoice numbers of the unconfirmed accounts receivable to Mr. Sawyer. He mentioned that he thought you had finished testing AR a few weeks ago. He has provided you the shipping documents for the accounts which he manages. He informed you that Kate Duke manages the remaining accounts and Meg Black, the Accounts Receivable supervisor, keeps the invoices until they are paid.

Please click on the invoice number below to access the corresponding shipping document.

41756
41768
41786
41900
41965
Cut-Off Testing

Kate Burke reiterated and expanded upon the explanation for the error that you found in the cut-off testing a couple of weeks ago. She explained that Mars Liquid Detergents has different terms for returning merchandise than the rest of Sprandel Inc.'s customers. The title of the merchandise changes hands (for Mars Liquid Detergents) when the carrier picks up the merchandise. For all of Sprandel's other customers, the title of the merchandise does not change hands until the merchandise is delivered by the carrier. The merchandise was picked up by the carrier on 12/30/08 but was not delivered to Sprandel Inc. until 1/2/09. By mistake, a credit memo was not issued until 1/2/09 because that is the process for other Sprandel Inc. customers. Kate has indicated that, if you think it's necessary, Meg Black will discuss adjusting for the error with your manager Bill Davis.

Ms. Burke has indicated that she will provide any further support you require for additional cut-off testing for the accounts she manages. She suggested that you request the support by credit memo number. She said she's pretty sure Meg previously supplied you with the cut-off detail.
Cut-Off Testing

Meg Black reiterated and expanded upon the explanation for the error that you found in the cut-off testing a couple of weeks ago. She explained that Mars Liquid Detergents has different terms for returning merchandise than the rest of Sprandel Inc.'s customers. The title of the merchandise changes hands (for Mars Liquid Detergents) when the carrier picks up the merchandise. For all of Sprandel's other customers, the title of the merchandise does not change hands until the merchandise is delivered by the carrier. The merchandise was picked up by the carrier on 12/30/08 but was not delivered to Sprandel Inc. until 1/2/09. By mistake, a credit memo was not issued until 1/2/09 because that is the process for other Sprandel Inc. customers. Meg has indicated that, if you feel it's necessary, she will discuss adjusting for the error with your manager Bill Davis.

Ms Black also mentioned that you already have the detail for any additional cut-off testing you may want to perform and suggested any further support related to additional credit memo cut-off testing should be obtained from the Account Manager.
EXHIBIT 25. Invoice Example
EXHIBIT 26. Shipping Document Example

<table>
<thead>
<tr>
<th>Date: 10/18/08</th>
<th>BILL OF LADING – SHORT FORM – NOT NEGOTIABLE</th>
<th>Page 1 of 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SHIP FROM:</td>
<td>Container No.:</td>
</tr>
<tr>
<td></td>
<td>Spreadbill Inc.</td>
<td>6754 5th Avenue Blvd.</td>
</tr>
<tr>
<td></td>
<td>SID No.: 4051</td>
<td>Carrier Name:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trailer number:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Serial number(s):</td>
</tr>
<tr>
<td></td>
<td>SHIP TO:</td>
<td>SCAC:</td>
</tr>
<tr>
<td></td>
<td>Ganone Coca Cola Inc.</td>
<td>3885 Commerce Drive</td>
</tr>
<tr>
<td></td>
<td>CID No.: 4223</td>
<td></td>
</tr>
<tr>
<td></td>
<td>THIRD PARTY FREIGHT CHARGES BILL TO:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ganone Coca Cola Inc.</td>
<td>3885 Commerce Drive</td>
</tr>
<tr>
<td></td>
<td>CID No.: 4223</td>
<td></td>
</tr>
<tr>
<td>Special Instructions:</td>
<td>Freight Charge Terms: (must be charged at time of shipment only):</td>
<td></td>
</tr>
<tr>
<td>Customer Order Information:</td>
<td>Prepaid:</td>
<td>Collect:</td>
</tr>
<tr>
<td>Customer Order No.:</td>
<td></td>
<td>3rd Party:</td>
</tr>
<tr>
<td># of Packages:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additional Shipper Information:</td>
<td>24015</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>25,000 20 &amp; over able to move</td>
<td></td>
</tr>
</tbody>
</table>

Note: Liability for loss or damage in this shipment may be applicable. See 49 U.S.C. § 14706(a)(1), (a)(9), and (a)(12).

Received, subject to the condition that the receiver notifies the carrier in writing of any defects or damage within five days of receipt of the goods. Where the receiver does not notify the carrier in writing of any defects or damage within five days of receipt of the goods, the receiver shall be deemed to have accepted the goods in good condition. This document is negotiable according to the applicable regulations of the ICC.

Shipper's Signature/Date: [Signature], 10/18/08

Trailer Loaded: [Signature] By shippers
[Signature] By driver

Freight Counted: [Signature] By shippers
[Signature] By driver

Carrier's Signature/Date: [Signature], 10/18/08

Certified that all packages are received by the carrier in good order, except as noted.

172
EXHIBIT 27. Customer Correspondence Example

John Sawyer

Please use 2 to 3 words below to describe one issue you would like to discuss with John Sawyer, Account Manager. For example, if you wanted to request support for testing cash accounts (as opposed to AR), you might type “cash testing”, “bank reconciliations” or “bank statements”. You will be able to ask him about more than one item. However, please inquire about one item at a time.

Please press “enter” after typing your words.

> press all correspondence

Back to Main Menu
Penns Oil

Please click "continue" below to verify that you have asked Mr. Sawyer whether he has corresponded with Penns Oil regarding payment for invoice #41756 and/or invoice #41768.

Continue

If you are trying to ask about something else, please click below to try again.

Try Again
Mr. Sawyer has indicated that he has just received notice of an electronic payment from Penns Oil for invoice #41756 and invoice #41768. He indicated that you can ask internal audit to assist you in verifying receipt of payment, if necessary.

If you would like to verify receipt of payment with internal audit, click below

- **Internal Audit**

- **Ask Mr. Sawyer Another Question**

- **Back to Main Menu**
Internal Audit

Penns Oil

With the assistance of internal audit, you have verified the electronic receipt of payment from Penns Oil for invoice #41,756 and invoice #41,768 in the amount of $169,834.20. Based on review of the evidence with internal audit and the date of payment, you are comfortable that the amount was a valid receivable for Speudel Inc. as of 12/31/08.
Quit Confirmation Screen

If you are finished with the task, please click below to access a few short background questions and questions about the case in which you just participated. Once you click "finished," you will no longer be able to return to the case.

Finished

If you are not finished with the task, please click below to return to the main menu

Return
You have completed the simulation portion of the study.

At this point, please save the workpapers (Excel file) and close them. They should automatically save to the thumb drive. Please do not make any changes to the workpapers at this point.

On the next several screens, you will be asked to respond to some brief questions regarding the case, as well as questions relating to your own experiences.

Please click "continue" below to proceed.

Continue
This question relates to the simulated case you just completed.

In the case you just completed, how long did the manager take to return the workpapers to you?

- 2 days
- 14 days
This question relates to the simulated case you just completed.

In the case you just completed, do you feel the length of time the audit manager had the workpapers before returning them to you was long or short?

1. Very Short
2. Not Short
3. Very Short
4. Not Long
5. Very Short
6. Not Long
7. Very Long
8. Not Long
9. Very Long
10. Very Long
11. Very Long
This question relates to the simulated case you just completed.

In the case you just completed, how long did you expect the audit manager to take to return the workpapers to you (i.e., what was the anticipated turnaround time for the review)?

- 2 days
- 14 days
This question relates to the simulated case you just completed.

Consider the length of time the audit manager took to review the workpapers you just completed.

How much of a *priority* do you feel these workpapers were to the manager?

1  2  3  4  5  6  7  8  9  10  11

Low  Priority

High  Priority
This question relates to the simulated case you just completed.

Consider the length of time the audit manager took to review the workpapers you just completed.

How much of a control do you believe the audit manager had over the length of time he took to review the workpapers?

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

No

Control

Significant

Control
This question relates to the simulated case you just completed.

How much do you feel that the length of time the manager took to review the workpapers was due to factors related to you, personally (as opposed to factors related to the manager and/or audit environment)?

1. Not At All
2. Not Related
3. A Bit
4. Somewhat Related
5. Related
6. Very Much Related

To Me
This question relates to the simulated case you just completed.

How much do you feel that the length of time the manager took to review the workpapers was due to factors related to the manager and/or audit environment (e.g., manager’s time management skills, pressures associated with manager’s workload or the time of the year, manager’s prioritization of this workpaper)?

1 2 3 4 5 6 7 8 9 10 11

Not at All
Related to Manager’s Environment

Very Much
Related to Manager’s Environment
This question relates to the simulated case you just completed.

Consider the length of time the audit manager took to review the workpapers you just completed.

To what extent did the length of review time affect your perceptions of the manager’s authority over you/the audit (i.e., their influence/power as an authority figure on the audit)?

1 2 3 4 5 6 7 8 9 10 11

Strongly
Disagree

Did Not
Impact

Strongly
Agree

Year of
Authority

Year of
Authority
This question relates to the simulated case you just completed.

Consider the length of time the audit manager took to review the workpapers you just completed.

To what extent were you frustrated by the length of his review time?

1. Not at all
2. Slightly
3. Moderately
4. Somewhat
5. Very

on a scale of

Frustrated

Frustrated
This question relates to the simulated case you just completed.

Consider the length of time the audit manager took to review the workpapers you just completed.

To what extent did you feel *angry* due to the length of his review time?

1 2 3 4 5 6 7 8 9 10 11

Not at All

Angry

Very

Angry
These questions relate to the simulated case you just completed. Please answer both questions, then click "continue" below.

Consider the way in which the review notes were worded in the case you just completed:

In general, how concerned was the reviewer about the documentation in the workpapers?

Not Concerned

Very Concerned

In general, how concerned was the reviewer about reaching the appropriate audit conclusions?

Not Concerned

Very Concerned

Note: You need to answer both questions before you will be able to continue.
This question relates to the simulated case you just completed.

Consider the way in which the review notes were worded in the case you just completed.

How concerned do you believe the audit manager was that there may be undiscovered issues or errors in Sprandel Inc.’s Accounts Receivable?

Not Concerned

Very Concerned
This question relates to your own background or professional experiences.

If in practice you had experienced the circumstances in this case, do you believe you would have responded similarly to the way you responded here?

Not At All 2 3 4 5 6 7 8 9 10 11

Similarly Very
This question relates to your own background or professional experiences.

Please remember all of your responses are completely confidential.

What is your position at your firm?

- Staff
- Manager
- In Charge / Senior
- Partner
This question relates to your own background or professional experiences.

Based on your experience, how realistic were the 7 review notes presented in this case?

Not At All

1 2 3 4 5 6 7

Realistic

Very

Realistic
This question relates to your own background or professional experiences.

The likelihood that I could be assigned to an audit engagement like Sprandel Inc. in the future is

Very
Low

Very
High
This question relates to your own background or professional experiences.

The amount of experience I have auditing manufacturing clients is

1  2  3  4  5  6  7  8  9  10  11

Very Low  Very High
How many months of audit experience do you have?

(click on the arrows to increase or decrease the number of months of experience)

0 Months

Continue
Based on your experience, how long would it typically take for a reviewer to review and return Accounts Receivable workpapers to you (i.e., what is the typical reviewer turnaround time of AR workpapers)?

0 Days

Continue
This question relates to your own background or professional experiences.

How much experience do you have closing review notes?

1 2 3 4 5 6 7 8 9 10 11

No Experience

Extensive Experience
This question relates to your own background or professional experiences.

Please indicate your gender:

- Female
- Male
This question relates to your own background or professional experiences.

How would you characterize your firm?

- International
- National
- Large Regional
- Small Regional
- Local
Feedback

If you would like to provide the researchers any feedback regarding this project (any thoughts regarding the simulation, any issues navigating through the simulation, or any other thoughts regarding this project, please use the space provided below. When you are done writing your comments, or if you do not have any to write, please press “finished” below to exit this simulation.
Table 1: Participants by Cell

<table>
<thead>
<tr>
<th></th>
<th>Conclusion Frame</th>
<th>Documentation Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timely Review</td>
<td>n = 16 (Cell 1)</td>
<td>n = 18 (Cell 2)</td>
</tr>
<tr>
<td>Delayed Review</td>
<td>n = 16 (Cell 3)</td>
<td>n = 19 (Cell 4)</td>
</tr>
</tbody>
</table>
Table 2: Sample Demographic Data

<table>
<thead>
<tr>
<th>Demographic Variable</th>
<th>Sample Mean</th>
<th>Standard Deviation</th>
<th>p-value&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Months of experience&lt;sup&gt;a&lt;/sup&gt;</td>
<td>16.68</td>
<td>13.63</td>
<td>0.420</td>
</tr>
<tr>
<td>Experience closing review notes&lt;sup&gt;b&lt;/sup&gt;</td>
<td>7.54</td>
<td>3.09</td>
<td>0.463</td>
</tr>
<tr>
<td>Experience with manufacturing clients&lt;sup&gt;c&lt;/sup&gt;</td>
<td>4.91</td>
<td>3.11</td>
<td>0.695</td>
</tr>
<tr>
<td>Likelihood of being assigned to similar client&lt;sup&gt;d&lt;/sup&gt;</td>
<td>8.23</td>
<td>2.47</td>
<td>0.685</td>
</tr>
</tbody>
</table>

<sup>a</sup> Audit experience measured in months.
<sup>b</sup> Participants responded on an 11-point scale (1 = “no experience” and 11 = “extensive experience”).
<sup>c</sup> Participants responded on an 11-point scale (1 = “very low” and 11 = “very high”).
<sup>d</sup> Participants responded on a 11-point scale (1 = “very low” and 11 = “very high”).
<sup>e</sup> p-value presented is from an ANOVA using the demographic variable as the dependent variable, as there are no directional expectations p-value is two-tailed.

<table>
<thead>
<tr>
<th>n</th>
<th>Percent</th>
<th>p-value&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff Auditor</td>
<td>53</td>
<td>77</td>
</tr>
<tr>
<td>Senior Auditor</td>
<td>16</td>
<td>23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>n</th>
<th>Percent</th>
<th>p-value&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>44</td>
<td>64</td>
</tr>
<tr>
<td>Female</td>
<td>25</td>
<td>36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>n</th>
<th>Percent</th>
<th>p-value&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>International Firm</td>
<td>65</td>
<td>94</td>
</tr>
<tr>
<td>National Firm</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Large Regional Firm</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 3: Interpretation of Underlying Directive (Pre-Test One)

<table>
<thead>
<tr>
<th>Review Note Set</th>
<th>Same/Similar Underlying Directive</th>
<th>Proportion</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>31 out of 33</td>
<td>93.9</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>30 out of 31</td>
<td>96.8</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>29 out of 31</td>
<td>93.5</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>29 out of 31</td>
<td>93.5</td>
<td></td>
</tr>
</tbody>
</table>

*Participants responded to the question: “How similar are the underlying directives (the ‘what to do’)?”*
Table 4: Perceptions of Frame Manipulation (Pre-Test One)

<table>
<thead>
<tr>
<th>Review Note</th>
<th>Mean Difference&lt;sup&gt;a&lt;/sup&gt;</th>
<th>N</th>
<th>Std. Deviation</th>
<th>t-stat&lt;sup&gt;b&lt;/sup&gt;</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Note A Documentation Frame</td>
<td>4.73</td>
<td>33</td>
<td>3.745</td>
<td>7.253</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Review Note A Conclusion Frame</td>
<td>3.82</td>
<td>33</td>
<td>4.004</td>
<td>5.479</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Review Note B Documentation Frame</td>
<td>4.85</td>
<td>33</td>
<td>3.124</td>
<td>8.916</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Review Note B Conclusion Frame</td>
<td>4.97</td>
<td>33</td>
<td>3.618</td>
<td>7.890</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Review Note C Documentation Frame</td>
<td>4.71</td>
<td>31</td>
<td>3.708</td>
<td>7.073</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Review Note C Conclusion Frame</td>
<td>3.36</td>
<td>31</td>
<td>3.738</td>
<td>4.998</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Review Note D Documentation Frame</td>
<td>3.03</td>
<td>32</td>
<td>4.036</td>
<td>4.249</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Review Note D Conclusion Frame</td>
<td>3.91</td>
<td>32</td>
<td>3.236</td>
<td>6.828</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

<sup>a</sup> Participants responded to the questions: (1) “How concerned was the reviewer about the documentation in the workpapers?” and (2) “How concerned was the reviewer about the appropriate audit conclusion?” (0=not concerned, 10=very concerned). The mean difference presented for documentation-framed review notes subtracts the response to question (2) from the response to question (1). The mean difference presented for conclusion-framed review notes subtracts the response to question (1) from the response to question (2).

<sup>b</sup> t-statistic and one-tailed p-value are based on one sample t-test comparisons to 0.
Table 5: Pretest Two: Analysis of Reviewer Expectation by Frame

<table>
<thead>
<tr>
<th>Review Note</th>
<th>Question</th>
<th>Frame</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>t-stat</th>
<th>p-value&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Note 3</td>
<td>a&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Conclusion</td>
<td>9</td>
<td>3.56</td>
<td>.726</td>
<td>0.140</td>
<td>0.890</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation</td>
<td>10</td>
<td>3.50</td>
<td>.972</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Conclusion</td>
<td>9</td>
<td>3.56</td>
<td>.527</td>
<td>-0.136</td>
<td>0.894</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation</td>
<td>10</td>
<td>3.60</td>
<td>.843</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Conclusion</td>
<td>9</td>
<td>4.11</td>
<td>.601</td>
<td>0.334</td>
<td>0.742</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation</td>
<td>10</td>
<td>4.00</td>
<td>.816</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review Note 5</td>
<td>a</td>
<td>Conclusion</td>
<td>9</td>
<td>3.89</td>
<td>.601</td>
<td>0.982</td>
<td>0.342</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation</td>
<td>10</td>
<td>3.50</td>
<td>1.080</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>Conclusion</td>
<td>9</td>
<td>3.67</td>
<td>.500</td>
<td>0.910</td>
<td>0.380</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation</td>
<td>10</td>
<td>3.30</td>
<td>1.160</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>Conclusion</td>
<td>9</td>
<td>4.33</td>
<td>.500</td>
<td>0.434</td>
<td>0.670</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation</td>
<td>10</td>
<td>4.20</td>
<td>.789</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review Note 6</td>
<td>a</td>
<td>Conclusion</td>
<td>9</td>
<td>3.28</td>
<td>1.149</td>
<td>0.503</td>
<td>0.621</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation</td>
<td>10</td>
<td>3.00</td>
<td>1.247</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>Conclusion</td>
<td>9</td>
<td>3.22</td>
<td>1.093</td>
<td>0.231</td>
<td>0.820</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation</td>
<td>10</td>
<td>3.10</td>
<td>1.197</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>Conclusion</td>
<td>9</td>
<td>3.78</td>
<td>.833</td>
<td>-0.510</td>
<td>0.960</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation</td>
<td>10</td>
<td>3.80</td>
<td>1.033</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review Note 7</td>
<td>a</td>
<td>Conclusion</td>
<td>8</td>
<td>4.13</td>
<td>.641</td>
<td>0.402</td>
<td>0.693</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation</td>
<td>10</td>
<td>4.00</td>
<td>.667</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>Conclusion</td>
<td>9</td>
<td>3.89</td>
<td>.782</td>
<td>-0.302</td>
<td>0.766</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation</td>
<td>10</td>
<td>4.00</td>
<td>.816</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>Conclusion</td>
<td>9</td>
<td>4.33</td>
<td>.500</td>
<td>0.434</td>
<td>0.670</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Documentation</td>
<td>10</td>
<td>4.20</td>
<td>.789</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Participants responded to the question: “Given the wording of this review note, how thoroughly do you expect the staff member to perform the procedure(s) you are requesting?” (1 = not at all thoroughly, 5 = very thoroughly).

<sup>b</sup> Participants responded to the question: “Suppose there is an issue or error [in the area being tested by each specific review note] which was not discovered in the additional round of testing. Please indicate your agreement with the following statement: In responding to this review note, I expect that a staff member would perform the procedures requested in such a way that he/she should find the problem.” (1 = strongly disagree, 5 = strongly agree).

<sup>c</sup> Participants responded to the question: “Given the wording of this review note, how important do you feel it is that the preparer performs the requested procedures? (1 = not at all important, 5 = very important)

<sup>d</sup> t-statistic and two-tailed p-value are based on independent sample t tests.
Table 6: Timeliness-Related Post-Experimental Questions

Panel A: Independent samples t-test results for questions in which there are directional expectations related to manipulation

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Timely Review</th>
<th>Delayed Review</th>
<th>t-stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of time to review was long or short(^a)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior to manager(^b)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perception of manager’s authority(^c)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frustrated due to length of time(^d)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angry due to length of time(^e)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Independent t-test results for questions in which there are no directional expectations related to manipulation

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Timely Review</th>
<th>Delayed Review</th>
<th>t-stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manager control over length of time to review(^f)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length of review time related to self(^g)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length of review time related to manager/environment(^h)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Participants responded on an 11-point scale (1 = very short, 11 = very long).
\(^b\) Participants responded on an 11-point scale (1 = low priority, 11 = high priority).
\(^c\) Participants responded on an 11-point scale (1 = strongly decreased view of authority, 6 = did not impact, 11 = strongly increased view of authority).
\(^d\) Participants responded on an 11-point scale (1 = not at all frustrated, 11 = very frustrated).
\(^e\) Participants responded on an 11-point scale (1 = not at all angry, 11 = very angry).
\(^f\) Participants responded on an 11-point scale (1 = no control, 11 = significant control).
\(^g\) Participants responded on an 11-point scale (1 = not at all related to me, 11 = very much related to me).
\(^h\) Participants responded on an 11-point scale (1 = not at all related to manager/environment, 11 = very much related to manager/environment).
\(^i\) Due to directional expectations related to the manipulation, p-values are one-tailed.
\(^j\) As there are no directional expectations, p-values are two-tailed.
Table 7: Frame-Related Post-Experimental Questions

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Conclusion Frame</th>
<th>Documentation Frame</th>
<th>t-stat</th>
<th>p-value^c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reviewer concern for documentation(^a)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>6.47 (2.51)</td>
<td>7.84 (2.42)</td>
<td>-2.301</td>
<td>0.013</td>
</tr>
<tr>
<td>Reviewer concern for conclusions(^a)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>8.44 (2.03)</td>
<td>7.95 (2.82)</td>
<td>0.839</td>
<td>0.202</td>
</tr>
<tr>
<td>Reviewer concern for unexpected issues/errors(^a)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>7.53 (2.08)</td>
<td>7.03 (2.59)</td>
<td>0.897</td>
<td>0.187</td>
</tr>
<tr>
<td>Relative concern for conclusions(^b)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>1.97 (2.73)</td>
<td>0.11 (2.96)</td>
<td>2.699</td>
<td>0.009</td>
</tr>
</tbody>
</table>

\(^a\) Participants responded on an 11-point scale (1 = not concerned, 11 = very concerned).
\(^b\) Participant response to reviewer concern for conclusions minus participant response to reviewer concern for documentation.
\(^c\) Due to directional expectations related to the manipulation, p-values are one-tailed.
Table 8: Number of Evidence Items Examined (Results of Hypotheses Testing)

Panel A: ANOVA results for number of relevant evidence items examined by the preparer

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>df</th>
<th>F-Statistic</th>
<th>p-value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Note Frame</td>
<td>1</td>
<td>2.237</td>
<td>0.066</td>
</tr>
<tr>
<td>Review Timeliness</td>
<td>1</td>
<td>2.818</td>
<td>0.049</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>0.762</td>
<td>0.193</td>
</tr>
</tbody>
</table>

Panel B: Mean (standard deviation) number of relevant evidence items examined by the preparer

<table>
<thead>
<tr>
<th>Timely Review</th>
<th>Conclusion Frame</th>
<th>Documentation Frame</th>
<th>Row Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cell 1</td>
<td>Cell 2</td>
<td>(SD)</td>
</tr>
<tr>
<td>Mean</td>
<td>31.38</td>
<td>25.50</td>
<td>28.26</td>
</tr>
<tr>
<td>(SD)</td>
<td>(9.55)</td>
<td>(9.86)</td>
<td>(10.02)</td>
</tr>
<tr>
<td>Delayed Review</td>
<td>Cell 3</td>
<td>Cell 4</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>25.13</td>
<td>23.53</td>
<td>24.26</td>
</tr>
<tr>
<td>(SD)</td>
<td>(10.42)</td>
<td>(10.65)</td>
<td>(10.42)</td>
</tr>
<tr>
<td>Column Means</td>
<td>28.25</td>
<td>24.49</td>
<td></td>
</tr>
<tr>
<td>(SD)</td>
<td>(10.33)</td>
<td>(10.18)</td>
<td></td>
</tr>
</tbody>
</table>

Panel C: Contrast tests between groups for number of relevant evidence items examined by the preparer

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Estimate</th>
<th>t-statistic</th>
<th>p-value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell 1 vs. Cell 2</td>
<td>5.88</td>
<td>1.685</td>
<td>0.049</td>
</tr>
<tr>
<td>Cell 1 vs. Cell 3</td>
<td>6.25</td>
<td>1.742</td>
<td>0.043</td>
</tr>
<tr>
<td>Cell 1 vs. Cell 4</td>
<td>7.85</td>
<td>2.280</td>
<td>0.013</td>
</tr>
<tr>
<td>Cell 2 vs. Cell 3</td>
<td>0.38</td>
<td>0.108</td>
<td>0.458</td>
</tr>
<tr>
<td>Cell 2 vs. Cell 4</td>
<td>1.97</td>
<td>0.591</td>
<td>0.278</td>
</tr>
<tr>
<td>Cell 3 vs. Cell 4</td>
<td>1.60</td>
<td>0.448</td>
<td>0.322</td>
</tr>
</tbody>
</table>

<sup>a</sup> Because directional effects are expected for all analyses, all p-values are one-tailed.
Table 9: Time Spent Closing Review Notes (Results of Hypotheses Testing)

Panel A: ANOVA results for time spent closing review notes

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>df</th>
<th>F-Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Timeliness</td>
<td>1</td>
<td>2.308</td>
<td>0.067</td>
</tr>
<tr>
<td>Review Note Frame</td>
<td>1</td>
<td>0.208</td>
<td>0.325</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>1.129</td>
<td>0.146</td>
</tr>
</tbody>
</table>

Panel B: Mean (standard deviation) amount of time spent closing review notes

<table>
<thead>
<tr>
<th>Timely Review</th>
<th>Conclusion Frame</th>
<th>Documentation Frame</th>
<th>Row Means (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cell 1</td>
<td>Cell 2</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>90.20 (32.68)</td>
<td>76.26 (33.56)</td>
<td>82.82 (33.40)</td>
</tr>
<tr>
<td>Delayed</td>
<td>Cell 3</td>
<td>Cell 4</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>70.16 (34.00)</td>
<td>72.84 (39.84)</td>
<td>71.61 (36.77)</td>
</tr>
</tbody>
</table>

Column Means (SD) | 80.18 (34.34) | 74.50 (36.44) |

Panel C: Contrast tests between groups for mean amount of time spent closing review notes

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Estimate</th>
<th>t-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell 1 vs. Cell 2</td>
<td>13.94</td>
<td>1.149</td>
<td>0.128</td>
</tr>
<tr>
<td>Cell 1 vs. Cell 3</td>
<td>20.04</td>
<td>1.605</td>
<td>0.057</td>
</tr>
<tr>
<td>Cell 1 vs. Cell 4</td>
<td>17.36</td>
<td>1.449</td>
<td>0.076</td>
</tr>
<tr>
<td>Cell 2 vs. Cell 3</td>
<td>6.10</td>
<td>0.503</td>
<td>0.609</td>
</tr>
<tr>
<td>Cell 2 vs. Cell 4</td>
<td>3.42</td>
<td>0.294</td>
<td>0.885</td>
</tr>
<tr>
<td>Cell 3 vs. Cell 4</td>
<td>-2.68</td>
<td>-0.224</td>
<td>0.885</td>
</tr>
</tbody>
</table>

a Time spent closing review notes is adjusted for pauses greater than 20 minutes.
b Because directional effects are expected for all analyses, all p-values are one-tailed.
Table 10: Time Spent Closing Review Notes (Robustness Test One)

Panel A: ANOVA results for time spent closing review notes

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>df</th>
<th>F-Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Note Frame</td>
<td>1</td>
<td>0.237</td>
<td>0.314</td>
</tr>
<tr>
<td>Review Timeliness</td>
<td>1</td>
<td>2.514</td>
<td>0.059</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>1.106</td>
<td>0.149</td>
</tr>
</tbody>
</table>

Panel B: Mean (standard deviation) amount of time spent closing review notes

<table>
<thead>
<tr>
<th>Timely Review</th>
<th>Conclusion Frame</th>
<th>Documentation Frame</th>
<th>Row Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cell 1</td>
<td>Cell 2</td>
<td>(SD)</td>
</tr>
<tr>
<td>Mean</td>
<td>86.36</td>
<td>74.41</td>
<td>80.03</td>
</tr>
<tr>
<td>(SD)</td>
<td>(29.131)</td>
<td>(32.00)</td>
<td>(30.82)</td>
</tr>
<tr>
<td>Delayed Review</td>
<td>Cell 3</td>
<td>Cell 4</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>65.88</td>
<td>70.26</td>
<td>68.26</td>
</tr>
<tr>
<td>(SD)</td>
<td>(30.41)</td>
<td>(35.92)</td>
<td>(33.10)</td>
</tr>
</tbody>
</table>

Column Means (SD) 76.12 (31.08) 72.28 (33.66)

Panel C: Contrast tests between cells for mean amount of time spent closing review notes

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Estimate</th>
<th>t-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell 1 vs. Cell 2</td>
<td>11.95</td>
<td>1.081</td>
<td>0.142</td>
</tr>
<tr>
<td>Cell 1 vs. Cell 3</td>
<td>20.48</td>
<td>1.801</td>
<td>0.038</td>
</tr>
<tr>
<td>Cell 1 vs. Cell 4</td>
<td>16.10</td>
<td>1.475</td>
<td>0.073</td>
</tr>
<tr>
<td>Cell 2 vs. Cell 3</td>
<td>8.53</td>
<td>0.772</td>
<td>0.222</td>
</tr>
<tr>
<td>Cell 2 vs. Cell 4</td>
<td>4.15</td>
<td>0.592</td>
<td>0.348</td>
</tr>
<tr>
<td>Cell 3 vs. Cell 4</td>
<td>-4.39</td>
<td>-0.402</td>
<td>0.345</td>
</tr>
</tbody>
</table>

*a Time spent closing review notes is adjusted for pauses greater than ten minutes.

b Because directional effects are predicted for all analyses, all p-values are one-tailed.
### Table 11: Time Spent Closing Review Notes (Robustness Test Two)

#### Panel A: ANOVA results for ranked time spent closing review notes

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>df</th>
<th>F-Statistic</th>
<th>p-value$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Note Frame</td>
<td>1</td>
<td>1.260</td>
<td>0.133</td>
</tr>
<tr>
<td>Review Timeliness</td>
<td>1</td>
<td>2.038</td>
<td>0.079</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>0.265</td>
<td>0.604</td>
</tr>
</tbody>
</table>

#### Panel B: Mean (standard deviation) ranked time spent closing review notes

<table>
<thead>
<tr>
<th>Timely Review</th>
<th>Conclusion Frame</th>
<th>Documentation Frame</th>
<th>Row Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cell 1</td>
<td>Cell 2</td>
<td>38.44 (18.40)</td>
</tr>
<tr>
<td>Mean</td>
<td>42.63</td>
<td>34.72</td>
<td>38.44</td>
</tr>
<tr>
<td>(SD)</td>
<td>(18.36)</td>
<td>(18.14)</td>
<td>(18.40)</td>
</tr>
<tr>
<td>Delayed Review</td>
<td>Cell 3</td>
<td>Cell 4</td>
<td>31.66 (21.28)</td>
</tr>
<tr>
<td>Mean</td>
<td>33.25</td>
<td>30.32</td>
<td>31.66</td>
</tr>
<tr>
<td>(SD)</td>
<td>(21.73)</td>
<td>(21.40)</td>
<td>(21.28)</td>
</tr>
<tr>
<td>Column Means (SD)</td>
<td>37.94 (20.35)</td>
<td>32.46 (19.73)</td>
<td></td>
</tr>
</tbody>
</table>

#### Panel C: Contrast tests between cells for ranked time spent closing review notes

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Estimate</th>
<th>t-statistic</th>
<th>p-value$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell 1 vs. Cell 2</td>
<td>7.90</td>
<td>1.151</td>
<td>0.127</td>
</tr>
<tr>
<td>Cell 1 vs. Cell 3</td>
<td>9.38</td>
<td>1.327</td>
<td>0.095</td>
</tr>
<tr>
<td>Cell 1 vs. Cell 4</td>
<td>12.31</td>
<td>1.815</td>
<td>0.037</td>
</tr>
<tr>
<td>Cell 2 vs. Cell 3</td>
<td>1.47</td>
<td>0.214</td>
<td>0.416</td>
</tr>
<tr>
<td>Cell 2 vs. Cell 4</td>
<td>4.41</td>
<td>0.670</td>
<td>0.253</td>
</tr>
<tr>
<td>Cell 3 vs. Cell 4</td>
<td>2.93</td>
<td>0.433</td>
<td>0.334</td>
</tr>
</tbody>
</table>

$^a$ Participants’ time spent closing review notes was ranked and recoded by the rank ordering of the data (e.g., lowest time is recorded as 1, second lowest is recorded as 2, and so on).

$^b$ Because directional effects are expected for all analyses, all p-values are one-tailed.
Table 12: Total Errors Found (Results of Hypotheses Testing)

Panel A: ANOVA results for total errors found

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>df</th>
<th>F-Statistic</th>
<th>p-value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Note Frame</td>
<td>1</td>
<td>0.348</td>
<td>0.279</td>
</tr>
<tr>
<td>Review Timeliness</td>
<td>1</td>
<td>2.301</td>
<td>0.067</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>2.399</td>
<td>0.063</td>
</tr>
</tbody>
</table>

Panel B: Mean (standard deviation) errors found

<table>
<thead>
<tr>
<th></th>
<th>Conclusion Frame</th>
<th>Documentation Frame</th>
<th>Row Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timely Review</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>Cell 1</td>
<td>Cell 2</td>
<td>(SD) 3.15</td>
</tr>
<tr>
<td>(SD)</td>
<td>3.56</td>
<td>2.78</td>
<td>(1.67)</td>
</tr>
<tr>
<td>Delayed Review</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>Cell 3</td>
<td>Cell 4</td>
<td>(1.56)</td>
</tr>
<tr>
<td>(SD)</td>
<td>2.44</td>
<td>2.79</td>
<td></td>
</tr>
</tbody>
</table>

Column Means (SD) 3.00 (1.50) 2.78 (1.58)

Panel C: Contrast tests between groups for total errors found

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Estimate</th>
<th>t-statistic</th>
<th>p-value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell 1 vs. Cell 2</td>
<td>0.78</td>
<td>1.484</td>
<td>0.072</td>
</tr>
<tr>
<td>Cell 1 vs. Cell 3</td>
<td>0.97</td>
<td>1.817</td>
<td>0.037</td>
</tr>
<tr>
<td>Cell 1 vs. Cell 4</td>
<td>0.77</td>
<td>1.480</td>
<td>0.072</td>
</tr>
<tr>
<td>Cell 2 vs. Cell 3</td>
<td>0.19</td>
<td>0.364</td>
<td>0.359</td>
</tr>
<tr>
<td>Cell 2 vs. Cell 4</td>
<td>-0.01</td>
<td>-0.023</td>
<td>0.491</td>
</tr>
<tr>
<td>Cell 3 vs. Cell 4</td>
<td>-0.20</td>
<td>-0.375</td>
<td>0.349</td>
</tr>
</tbody>
</table>

<sup>a</sup> Because directional effects are expected for all analyses, all p-values are one-tailed.
Table 13: Over-documentation of Procedures (Measures One and Two)

Panel A: ANOVA results for total instances of over-documentation (measure one)\(^a\)

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>df</th>
<th>F-Statistic</th>
<th>p-value(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Note Frame</td>
<td>1</td>
<td>0.024</td>
<td>0.439</td>
</tr>
<tr>
<td>Review Timeliness</td>
<td>1</td>
<td>1.286</td>
<td>0.131</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>0.111</td>
<td>0.370</td>
</tr>
</tbody>
</table>

Panel B: Mean (standard deviation) instances of over-documentation (measure one)\(^a\)

<table>
<thead>
<tr>
<th>Timely Review</th>
<th>Conclusion Frame</th>
<th>Documentation Frame</th>
<th>Row Means (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Cell 1</td>
<td>Cell 2</td>
<td>1.91 (3.76)</td>
</tr>
<tr>
<td>(SD)</td>
<td>1.63 (3.50)</td>
<td>2.17 (4.06)</td>
<td></td>
</tr>
<tr>
<td>Delayed Review</td>
<td>Cell 3</td>
<td>Cell 4</td>
<td>3.14 (5.15)</td>
</tr>
<tr>
<td>Mean</td>
<td>3.25 (4.87)</td>
<td>3.05 (5.51)</td>
<td></td>
</tr>
<tr>
<td>(SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Column Means (SD) 2.44 (4.25) 2.62 (4.82)

Panel C: ANOVA results for proportion of participants who over-documented (measure two)\(^c\)

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>df</th>
<th>F-Statistic</th>
<th>p-value(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Note Frame</td>
<td>1</td>
<td>0.101</td>
<td>0.376</td>
</tr>
<tr>
<td>Review Timeliness</td>
<td>1</td>
<td>1.338</td>
<td>0.126</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>0.006</td>
<td>0.469</td>
</tr>
</tbody>
</table>

Panel D: Mean (standard deviation) proportion of participants who over-documented (measure two)\(^c\)

<table>
<thead>
<tr>
<th>Timely Review</th>
<th>Conclusion Frame</th>
<th>Documentation Frame</th>
<th>Row Means (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Cell 1</td>
<td>Cell 2</td>
<td>0.26 (0.45)</td>
</tr>
<tr>
<td>(SD)</td>
<td>0.25 (0.45)</td>
<td>0.28 (0.46)</td>
<td></td>
</tr>
<tr>
<td>Delayed Review</td>
<td>Cell 3</td>
<td>Cell 4</td>
<td>0.40 (0.50)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.38 (0.50)</td>
<td>0.42 (0.51)</td>
<td></td>
</tr>
<tr>
<td>(SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Column Means (SD) 0.31 (0.47) 0.35 (0.48)

\(^a\) Variable is measured as total number of times a participant over-documents procedures.

\(^b\) Because directional effects are expected for all analyses, all p-values are one-tailed.

\(^c\) Variable is coded as 1 if participant over-documents any items and 0 if participant does not over-document at all.
Table 14: Over-documentation of Procedures (Measure Three)

Panel A: ANOVA results for percentage of over-documentation

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>df</th>
<th>F-Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Note Frame</td>
<td>1</td>
<td>0.101</td>
<td>0.316</td>
</tr>
<tr>
<td>Review Timeliness</td>
<td>1</td>
<td>1.338</td>
<td>0.043</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>0.006</td>
<td>0.158</td>
</tr>
</tbody>
</table>

Panel B: Mean (standard deviation) percentage of over-documentation

<table>
<thead>
<tr>
<th>Timely Review</th>
<th>Conclusion Frame</th>
<th>Documentation Frame</th>
<th>Row Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Cell 1</td>
<td>Cell 2</td>
<td>(SD)</td>
</tr>
<tr>
<td>(SD)</td>
<td>0.05</td>
<td>0.09</td>
<td>0.07</td>
</tr>
</tbody>
</table>

| Delayed Review              | Cell 3           | Cell 4              | (SD)      |
| Mean                        | 0.24             | 0.14                | 0.19      |
| (SD)                        | (0.41)           | (0.30)              | (0.35)    |

Column Means (SD) 0.15 (0.31) 0.11 (0.34)

Panel C: Contrast tests between groups for percentage of over-documentation

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Estimate</th>
<th>t-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell 1 vs. Cell 2</td>
<td>-0.35</td>
<td>-0.375</td>
<td>0.355</td>
</tr>
<tr>
<td>Cell 1 vs. Cell 3</td>
<td>-0.18</td>
<td>-1.898</td>
<td>0.031</td>
</tr>
<tr>
<td>Cell 1 vs. Cell 4</td>
<td>-0.08</td>
<td>-0.897</td>
<td>0.187</td>
</tr>
<tr>
<td>Cell 2 vs. Cell 3</td>
<td>-0.15</td>
<td>-1.578</td>
<td>0.060</td>
</tr>
<tr>
<td>Cell 2 vs. Cell 4</td>
<td>-0.05</td>
<td>-0.538</td>
<td>0.296</td>
</tr>
<tr>
<td>Cell 3 vs. Cell 4</td>
<td>0.10</td>
<td>1.056</td>
<td>0.217</td>
</tr>
</tbody>
</table>

a Variable is measured calculated as the cumulative count of the number of evidence items overdocumented divided by the total number of evidence items accessed by the participant.

b Because directional effects are expected for all analyses, all p-values are one-tailed.
VITA

TAMARA A. LAMBERT

EDUCATION

Doctor of Philosophy in Business Administration (Accounting)
Drexel University, Philadelphia, PA  September 2009

Bachelor of Science/Bachelor of Arts
Major Field: Accounting, Minor Field: English (magna cum laude)
Bloomsburg University, Bloomsburg, PA  May 2001

PROFESSIONAL EXPERIENCE

Academic
2009  Assistance Professor, University of Massachusetts Amherst
2004-2009  Graduate Teaching and Research Assistant, Drexel University

Business
2001-2004  Assurance and Advisory Services Senior, Deloitte & Touche

RESEARCH

Doctoral Dissertation:
“Closing Review Notes: The Effect of Reviewer Delay and Review Note Frame on Audit Workpaper Preparers’ Effort and Performance”

Working Papers:

“An Examination of the Audit Manager’s Role in the Underreporting of Time” with Christopher P. Agoglia and Richard C. Hatfield.

PROFESSIONAL MEMBERSHIPS

Certified Public Accountant (Pennsylvania, 2003)
American Accounting Association

HONORS

American Accounting Association Doctoral Consortium Fellow, 2007
Lebow College Outstanding Ph.D. Student Instructor Award, 2007-2008