A REEXAMINATION OF THE DILUTION OF AUDITOR MISSTATEMENT RISK ASSESSMENTS: An experimental study of the impact of client information type, workload, and PCAOB guidance on dilution.

Suzanne M. Perry

A Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

October, 2015

Don W. Finn, Committee Chair

Gary F. Peters, Committee Member

Jesse C. Robertson, Committee Member

Robert J. Pavur, Committee Member

Marilyn Wiley, Dean of the College of Business

Costas Tsatsoulis, Dean of the Robert B. Toulouse School of Graduate Studies

Perry, Suzanne M. *A Reexamination of the Dilution of Auditor Misstatement Risk Assessments: An Experimental Study of the Impact of Client Information Type, Workload, and PCAOB Guidance on Dilution*. Doctor of Philosophy (Accounting), December 2015, 65 pp., 20 tables, 2 figures, references, 41 titles.

Many external parties such as investors, creditors, and regulatory agencies, use a company's financial statements in their decision-making. In doing so, they rely on audit opinions on whether financial statements are fairly stated. However, evidence suggests that there are factors in the audit environment that influence auditor judgments. For example, nondiagnostic client information dilutes auditor judgments when compared to judgments based on diagnostic information alone, especially for less experienced auditors (Hackenbrack 1992; Hoffman and Patton 1997; Glover 1994; Shelton 1999). High time pressure conditions mitigate this effect by refocusing auditor attention toward relevant client information, therefore reducing the impact of nondiagnostic information (Glover 1994, 1997).

This research study examines other common audit environment factors to determine if they too influence audit judgment results. An online questionnaire of 149 auditors, CPAs and other accounting professionals indicate that the inclusion of nondiagnostic client information results in a significant change in auditor judgments. The direction of this change follows a theorized pattern; risk assessments that were initially high are reduced, while those that were initially low are increased. Significance was not consistently found for a workload and PCAOB effect on auditor judgment. However, a comparison of the absolute value of dilution effect means across conditions reveals some trending for the proposed unwanted effect of high workload, and the beneficial effect of PCAOB guidance.

These results have important implications for auditing research and practice. It extends previous archival research on workload effects and uses a unique questionnaire design to reexamine workload pressures in a behavioral setting. The results of hypothesis testing on workload pressure and PCAOB guidance, although lacking consistent statistical significance; exhibit trends that agree with proposed theoretical relationships. Tests on the effects of nondiagnostic information show strong statistical support for previous studies in the area of psychology and audit. This study's greatest contribution suggests that audit pressures do not produce equivalent effects on auditor judgment; time pressure improves audit judgment, while workload pressure does not (Glover 1994, 1997). These results can be explained by examining the relationship between stress and audit judgment performance (Choo 1995, Yerkes and Dodson 1908). Different types and different degrees of audit pressures may correspond to different levels of audit pressure. Low to moderate levels of audit pressure, such as the level of time pressure used in Glover's (1994, 1997) study improve audit performance. Higher audit pressures, such as high workload during an auditor's busy season, may lower audit performance.

TABLE of CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1

INTRODUCTION

Research Objective

Auditors play a vital role by assessing whether the client's financial statements are fairly stated after conducting an impartial, independent, examination of its records. As a result, financial statement users rely on the audit opinion for related investment and credit decisions. Because of the external user's reliance on auditor opinions, high audit performance is essential. However, audit performance and auditor judgment are often influenced by various environmental constraints in the audit review process. One example would be completing the audit within budget (Enhancing the Value of Auditor Reporting, IAASB Consultation Paper 2011, 6, para. 8) and is referred to as time pressure. Auditors who experience time pressure must complete an audit task within a specified amount of time, and being subject to this constraint may impact performance.

The purpose of this study is to examine the impact of nondiagnostic client information, workload, and PCAOB guidance on dilution. Dilution is calculated using an adaptation of the Glover (1994, 1997) accounts receivable task, which asks auditors to make misstatement risk judgments in response to fictitious client cases. Using an audit task that builds onto a previously validated instrument will add comparative support to the dilution literature.

This study has both theoretical and practical implications. It will contribute to theory and research by clarifying inconclusive evidence from prior audit research, and validating prior dilution effect theory.

Relevant Literature

Prior research evidence have suggested that time pressure is related to lower audit effectiveness (McDaniel 1990). Lower audit performance (Braun 2000), as a result of time pressure, has been associated with reduced audit performance on an inventory task (Low and Tan 2011). However, the time pressure effect can be decreased when certain precautions are followed. For example, when auditors are warned of a time constraint early in the planning stage, rather than immediately before the testing phase of the audit task, the time pressure effect can be mitigated (Low and Tan 2011). Additionally, the interaction of an early warning and audit instructions can further reduce the time pressure effect (Low and Tan 2011).

Unlike McDaniel (1990), Braun (2000) and Low and Tan (2011), Glover's (1997) results indicated that an increase in time pressure was associated with more accurate misstatement risk judgments, compared to judgments that are less time constrained. He found evidence to suggest that auditor judgments of accounts receivable misstatement risk, made during low time pressure, were impacted more by nondiagnostic information[1] than judgments made when there is greater time pressure to perform the audit. The addition of nondiagnostic client information to case materials resulted in a dilution effect, [2] risk assessments that were more moderate compared to the auditors' original estimate of the risk of material misstatement when using diagnostic client information only. However, auditors who experienced higher time pressure exhibited a re-prioritization towards diagnostic client information, which resulted in

[1] Nondiagnostic client information is information that is irrelevant to judgments. Specific examples include work papers and excerpts of other audit procedures. Dilution occurs when nondiagnostic information decreases high-risk judgments and increases low-risk judgments, thus resulting in more moderate judgments (Nisbett 1981).
[2] The dilution effect was first identified in the psychology literature (Nisbett et al. 1981). Hackenbrack (1992) examined the dilution effect in an audit setting. Findings indicated that auditors estimated a smaller change in a firm's level of fraud-risk when a combination of non-diagnostic and diagnostic information was included compared to estimations based on diagnostic information only.

accounts receivable audit risk judgments that were less diluted, therefore causing an improvement in audit performance.

Even though time pressure has a beneficial effect on auditor judgment, it is only one of many job pressures that an auditor must face. Organizational characteristics, a subcategory of job pressure, are also commonly encountered and include feedback, social influence and workload (DeZoort and Lord 1997). Workload (i.e., the number of clients being audited) often fluctuates throughout the year, causing auditors to frequently experience an increase in work during busy season periods (Sweeney and Summers 2002; Lopez and Peters 2011; 2012). Experimental research examining workload effects is important to understand how it may influence the audit, especially if a workload increase results in reduced performance.

If an increase in workload has an unwanted impact on audit performance, audit firms would want to know how to mitigate that effect. One way might be to provide auditors with audit guidance related to the appropriate steps used to conduct an audit. This could be accomplished by including brief excerpts of applicable audit standards (PCAOB, SEC, or others) into the audit plan. Examining the impact of increased applicable audit standard guidance on auditors' misstatement risk judgments, combined with increased workload, may provide evidence of the impact of job pressure, while providing a possible counteracting solution. Understanding the impact of increased workload may also help explain the contradictory results of prior time pressure audit research.

Motivation

Three main variables are used in this study, nondiagnostic client information, workload, and audit standard guidance.[3] Nondiagnostic client information dilutes auditor judgment, thus reducing auditor performance (Hackenbrack 1992; Hoffman and Patton 1997; Glover 1997; Shelton 1999; and Wood 2012). This effect was seen in prior research in a variety of audit tasks including, fraud-risk, misstatement risk, and going-concern assessments. These results suggest that nondiagnostic information may lower audit performance in multiple areas in an audit. Since it is common for auditors to encounter these situations in the performance of an audit, it is important to continue to investigate nondiagnostic information effects on audits. This study will further validate the impact that nondiagnostic information can have on auditor misstatement risk assessments (adapted from Glover 1994, 1997). Doing so will confirm the presence of the dilution effect prior to the impact of workload and audit standard guidance on auditor performance.

Secondly, there has been extensive experimental audit research on the influence of time pressure on audit performance with differing results (McDaniel 1990; Braun 2000; Margheim et al. 2005; Bowrin and King II 2010; Glover 1997). Unlike McDaniel (1990), Braun (2000), Margheim et al. (2005), and Bowrin and King II (2010); Glover (1997) found evidence to support that time pressure improved audit performance. One theory that might account for these differences is a theoretical relationship between stress and performance; stress that is increased up until a moderate level improves performance (Choo 1995). However, if stress exceeds a moderate amount, an opposite effect occurs and performance declines. Glover's

[3] This study will use excerpts of audit standards taken from the PCAOB, however other sources of audit standard guidance (SEC, AICPA, etc) may be used.

(1997) findings may indicate that time pressure equates to a level of stress that is lower on the Yerkes and Dodson (1908) curve than other research studies, therefore improving audit performance rather than reducing it.

More recently, two archival studies suggest that workload compression[4], another source of auditor stress, is associated with lower audit quality and unwanted audit outcomes (Lopez and Peters 2011; 2012). This may also indicate that workload induces stress that is beyond a moderate level and result in reduced audit performance. Additional experimental research is needed to fully understand this issue.

Prior literature suggests that high workload will result in a larger dilution effect (McGrath 1976; Yerkes and Dodson 1908). However, because workload and time pressure are closely related, time pressure was introduced across all conditions in order to improve audit performance as seen in the Glover (1994, 1997) studies.

Finally, if increased workload results in greater dilution, the addition of nondiagnostic client information would also result in misstatement risk judgments that are less accurate. Because auditors routinely experience increased workload, audit firms need a way to mitigate its impact. Researchers have suggested that decision aides, or audit guidance, could change or improve the impact of time pressure on audit performance (Braun 2000; Hoffman and Patton 1997). Therefore, the third factor is the use of PCAOB guidance as a possible decision aid, operationalized as excerpts from PCAOB standards. Audit standards will remind auditors of the appropriate way to audit accounts receivable. Theory suggests that auditors will compare how they conduct the audit to the guidelines dictated in the PCAOB excerpt, thus choosing to act in

[4] Lopez and Peters (2011; 2012) define workload compression as the grouping of audit-firm clients by their fiscal year-end date.

a manner that is similar to the PCAOB ideal (Nisbett et al. 1981). As a result, auditors will prefer client information that is the most relevant to the audit judgment, therefore reducing the dilutive effects of nondiagnostic client information and improving audit performance.

Findings validate prior research on the dilutive effects of nondiagnostic information on auditor judgments. The addition of nondiagnostic client information to audit cases results in secondary audit judgments that are lower for high risk cases and higher for low risk cases. While trends exist for the proposed effects of workload and PCAOB guidance in some client cases, statistical significance is not consistently found. This may be due to the small sample size or the demographic nature of the respondents.

The next chapter examines prior literature that is pertinent to this study, and begins by reviewing studies of audit performance. Subsequent subsections discuss relevant literature for each factor used in this study including studies of dilution, and workload.

CHAPTER 2

LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

Accounting Studies of Auditor Performance

Audit performance, how closely an auditor meets an audit objective (Bowrin and King II 2010), has been the focus of much research. Prior studies have examined factors that influence audit performance for many years. Not only has there been much variation in the type of factors that have been investigated, operationalization of the audit performance measure has varied. For example, audit performance has been measured as audit effectiveness, efficiency, consistency, processing accuracy and subsidiary task performance, and dilution effect (McDaniel 1990; Braun 2000; Margheim et al. 2005; Bowrin 2001; Bowrin and King II 2010; Glover 1997). The following section will review some of the more notable audit performance measures, prior to a discussion of the external factors that might influence it.

Audit Effectiveness as an Audit Performance Measure

Audit researchers have used audit performance measures, such as audit effectiveness, in time pressure audit research with mixed results. Audit effectiveness is likened to the "job performance" of the auditor, which is consistent with the conceptual definition of effectiveness in psychology literature (McDaniel 1990, 268, footnote 1). However, it has been operationalized in different ways by different audit researchers (Bowrin and King II 2010; McDaniel 1990; Braun 2000; Margheim et al. 2005), which has made comparison across studies more difficult.

McDaniel (1990) measured audit effectiveness as the multiplication of processing accuracy and sampling adequacy in the auditor's identification of errors related to finished goods valuation, finished goods completeness, inventory reserve valuation, and inventory reserve completeness. Audit effectiveness scores for each of the four tasks were summed to create a total audit effectiveness measure for each participant. In a related study, Braun (2000) compared the performance of auditors on a dominant task (recording evidence of potential inventory misstatements) with their performance on a subsidiary task (the identification of client characteristics that might indicate fraud). Audit performance on the dominant task was measured using processing accuracy, only one part of McDaniel's (1990) two-part audit effectiveness score. Greater time pressure impacts these two measures of audit performance differently (discussed below).

Bowrin and King II (2010) used a different audit effectiveness scale calculated as the difference between accurate responses versus inaccurate responses on a ratio-analysis audit task, and the sum of accurate responses on a compliance-testing audit task. Each of the three preceding studies (Bowrin and King II 2010; McDaniel 1990; Braun 2000) required auditors to participate in an audit task. Therefore, the audit effectiveness scores measured the accuracy of the auditors' responses in the performance of the task. Conversely, Margheim et al. (2005, 27) relied on the auditors' probability estimations of whether the "hypothetical senior auditor and staff auditor would 'be highly' effective" in a fictitious audit scenario.

Many of the aforementioned studies examined the impact of time pressure on audit effectiveness (Bowrin and King II 2010; McDaniel 1990; Braun 2000; Margheim et al. 2005). Results of these studies have been mixed, which may be due to the differences in the

operationalization of the audit effectiveness measure.[5] Detailed results of these studies will be discussed in the time pressure section of this chapter.

Dilution Effect

Another audit performance proxy is the psychology-based construct known as the dilution effect. A series of psychology studies conducted in 1981 examined dilution by investigating the impact of information type[6] on judgment (Nisbett et al.). In these studies participants were required to make predictions about a subject's level of electric shock tolerance and number of movies attended based on diagnostic and nondiagnostic information about the subject. Diagnostic information included facts that were directly relevant to judgment, while nondiagnostic information included facts that were irrelevant. Nisbett et al. (1981) discovered that including nondiagnostic information resulted in the dilution of a participant's shock tolerance and movie attendance estimates about that subject. Nisbett et al. (1981) also identified an incremental effect of nondiagnostic information on likelihood predictions that a target individual was a child abuser, indicating greater dilution as more nondiagnostic information was added.

Nisbett et al. (1981) suggested that the dilution effect might be the result of participants' use of similarity based judgments, a comparison of how similar a target is to its desired outcome (Kahneman and Tversky 1972; 1973). Participants perceived that the individual was more similar to an internally held stereotype or heuristic when considering

[5] Otley and Fakiolas (2000) proposed that the conflicting results seen in the reliance on accounting performance measures (RAPM) was due to differences in the experimental tasks and the operationalization of the variable.
[6] Diagnostic versus nondiagnostic information

diagnostic information only (Nisbett et al. 1981). This resulted in judgments that were relatively more “extreme” when compared to judgments based on nondiagnostic and diagnostic information (Nisbett et al. 1981, 256). Nondiagnostic information reduced participants’ perception of the similarity between the subject and an internally held stereotype. Therefore, when nondiagnostic information was included with diagnostic information the judgments of the participants became diluted (Nisbett et al. 1981).

Dilution Effect as an Audit Performance Measure

Auditors also react differently to diagnostic and nondiagnostic client information. Results of prior studies have suggested that the inclusion of nondiagnostic information (information that is not relevant to the judgment), results in audit judgments that are more diluted when compared to judgments made using only diagnostic information (Glover 1997; Hackenbrack 1992; Hoffman and Patton 1997; Shelton 1999). This is due to auditors focusing concurrently on irrelevant and relevant information when forming judgments, which lowers performance (Easterbrook 1959).

This dilution effect occurs when nondiagnostic client information is added to diagnostic client information and results in auditor judgments that move toward the center of the misstatement risk scale (Hackenbrack 1992; Glover 1997). If an auditor initially judged a client to have a high risk of material misstatement (based on a review of client diagnostic information only), dilution would occur if the auditor’s second judgment (based on a review of the same client diagnostic information plus nondiagnostic information) lowered the client’s misstatement risk. Conversely, if the initial judgment was of a low misstatement risk, dilution would occur if

nondiagnostic client information resulted in an increase of the auditor's misstatement risk judgment. Dilution has been used as an audit performance measure in prior literature such as Glover (1997), Hackenbrack (1992), Hoffman and Patton (1997), and Shelton (1999).

Hackenbrack (1992) tested the dilution effect in an experimental audit environment using two tasks, a case-study audit task and a rating audit task. The case-study audit task required auditors to review a fictitious case containing both nondiagnostic and diagnostic information. The case included information about a client's activities prior to a fraudulent event, information about the fraudulent event (diagnostic information about an increased fraud-risk or decreased fraud-risk) and work papers (nondiagnostic information that was favorable, neutral or unfavorable) from various dates after the event but within the same year. Auditors then estimated the change in their perception of the firm's fraud-risk after reading the year-long work papers when compared to their impression based on the pre-fraud information alone. The rating audit task required auditors to rate (on two separate scales) the fraud-risk impact of each of five diagnostic scenarios that heightened a firm's fraud-risk and each of five diagnostic scenarios that diminished a firm's fraud-risk.

Hackenbrack's (1992) findings indicated that auditors estimated a smaller change in a firm's level of fraud-risk when a combination of nondiagnostic and diagnostic information was included compared to estimations based on diagnostic information only. This effect was stronger for favorable and unfavorable nondiagnostic information, compared to neutral nondiagnostic information, when increased fraud-risk scenarios were included (Hackenbrack 1992).

Hoffman and Patton (1997) extended Hackenbrack's (1992) study by examining the impact of accountability on fraud-risk judgments, to determine if accountability was associated with an increase in the dilution effect (lower estimations of fraud risk) or with conservative audit judgments (higher estimations of fraud risk). Auditors were required to estimate a client's fraud-risk for each of six cases. All cases contained two relevant scenarios that were categorized as strong and positive, otherwise they differed based on amount and type of relevant (strong/weak, positive/negative) and irrelevant (favorable/unfavorable) information. Findings indicated that both accountable and non-accountable[7] auditors exhibited the dilution effect, displaying lower fraud-risk estimates for cases that included irrelevant information. Additionally, auditors who were held accountable did not have judgments that were more diluted, but did exhibit fraud-risk estimates that were more conservative. Therefore, accountability worked against the dilution effect, resulting in fraud-risk estimates that were higher.

The results from Hoffman and Patton (1997) and Hackenbrack (1992) support the dilution effect in an audit environment; however, they differ on two main points. First, Hoffman and Patton (1997) included tests of accountability, which Hackenbrack (1992) discussed in his concluding remarks but did not incorporate in his experimental design. Second, while the audit judgments in both studies are fraud-related, subtle differences exist. Hackenbrack (1992) used a complicated two phase experiment to test his hypotheses. Both audit tasks required participants to rate their reaction to the client information; how much their fraud-risk estimate

[7] Auditors in the non-accountable condition were told their judgments would remain anonymous, while those in the accountable condition were told that they might be chosen to defend their responses to a panel of experienced auditors.

changed, increased, or decreased after reviewing the experimental information. Hackenbrack (1992) then measured the dilution effect as the case study rating (based on diagnostic and nondiagnostic information) minus the corresponding rescaled score from the rating audit task (based on diagnostic information). If the calculated measure is negative then the dilution effect occurred (Hackenbrack 1992). Conversely, Hoffman and Patton (1997) used a single experimental task that required auditors to rate the likelihood of fraud for each of six cases. The dilution effect was then calculated as the difference in an auditor's fraud-risk estimate between a case that contained relevant information and a case that contained a combination of relevant and irrelevant information (Hoffman and Patton 1997). Unlike Hackenbrack (1992), Hoffman and Patton's (1997) calculation produced a dilution effect with a positive sign. Then the impact of accountability on dilution was examined by comparing the dilution effect scores for the accountable auditors with those of the non-accountable auditors (Hoffman and Patton 1997).

Both Hackenbrack (1992) and Hoffman and Patton (1997) used auditors with limited experience averaging no more than three and half years. By contrast, Shelton (1999) suggested that an auditor's experience level might play a vital role in how irrelevant information is processed. Two groups of auditors (experienced: managers and partners, less experienced: seniors) provided a going-concern assessment based on a case containing either relevant information or both relevant and irrelevant information (Shelton 1999). Unlike Hackenbrack (1992) and Hoffman and Patton (1997), participants only reviewed one case and provided one judgment. In Shelton's (1999) study, the dilution effect was not calculated as a within-subject variable as it was in prior research. Findings indicated that auditors used irrelevant information

differently based on experience level. The going-concern assessments of less experienced auditors provided evidence in support of the dilution effect. However, irrelevant information did not significantly change the assessments of more experienced auditors.

Additionally, Glover (1997) tested the presence of the dilution effect using a two phase experiment. Phase one instructed the auditors to judge the accounts receivable misstatement risk for eight short hypothetical cases. Each of these eight cases included diagnostic client information; information that had a direct influence on the misstatement risk judgment. In phase two, auditors were asked to make an accounts receivable misstatement risk judgment based on one long case. The long case contained the same diagnostic information as one of the shorter cases and additional nondiagnostic client information that did not directly relate to the misstatement risk judgment. The dilution effect was calculated as the difference in misstatement risk judgment between the long case (nondiagnostic and diagnostic) and its corresponding shorter case (diagnostic). How the difference was calculated depended on the type of misstatement risk (high or low) for the short case. Glover (1997) found that including nondiagnostic client information resulted in the dilution of an auditor's misstatement risk judgment.

Prior researchers have suggested that dilution is the result of similarity-based judgments, which compare the target that is being judged to its perceived outcome (Hackenbrack 1992; Nisbett et al. 1981; Glover 1997). In an audit task, this can be explained as a comparison of hypothetical client information to an auditor's image of a high risk client. If an auditor perceives that the hypothetical client is similar to an internally held image of a high risk client, then the auditor will rate the client higher on the fraud risk or misstatement risk scale.

Including nondiagnostic client information lessens the similarity in the mind of the auditor, therefore causing the dilution effect (Hackenbrack 1992; Glover 1997).

Time Pressure as an External Audit Pressure

External pressures such as time pressure may also affect audit performance. A commonly experienced pressure such as time pressure, has garnered much interest in the auditing literature. Time pressure is usually operationalized as a reduction of the time that the auditor has to complete the experimental task, similar to a reduction of the time to perform an audit. This reduction has been simulated as a decrease in the number of minutes available or a reduction in the percentage of time available (McDaniel 1990; Braun 2000; Margheim et al. 2005; Bowrin and King II 2010; Glover 1997). Many of these researchers have drawn from foundation level psychology literature such as Yerkes and Dodson (1908), Easterbrook (1959), and McGrath (1976) in order to predict the impact of time pressure on audit effectiveness outcomes.

Time Pressure and Audit Effectiveness

Time pressure impacts audit effectiveness in a variety of ways. Results from McDaniel (1990), Margheim et al. (2005) and Bowrin and King II (2010) suggest that time pressure is associated with lower audit effectiveness. Conversely, using an inventory audit task, Braun (2000) found no association between time pressure and audit effectiveness on his primary audit task, although high time pressure was associated with lower performance on his secondary audit task. Braun (2000) stated that the results were not due to the time allowed; but instead,

to how the auditors chose to use their time. Braun (2000) cited the Easterbrook (1959) Hypothesis[8] as an explanation for these findings, which posits that an aroused state result in a smaller range of cues used; a re-prioritization of attention away from peripheral cues, while maintaining focus on central cues (Easterbrook 1953, 1959). He explained that high time pressure resulted in an increase in concentration given to the primary experimental audit task, and less concentration given to the secondary experimental audit task (Braun 2000).

Bowrin and King II (2010) found evidence to suggest that increased time pressure was associated with a greater reduction in audit effectiveness for complex audit tasks, when compared to simpler audit tasks. Margheim et al. (2005) also found time pressure to reduce audit effectiveness when two levels of inexperienced auditors were studied. Audit seniors and audit staff both exhibited a lower expectation of audit effectiveness under high time budget pressure and high time deadline pressure. McDaniel (1990) also found that increased time pressure was associated with lower audit effectiveness, with a reduction in both processing accuracy, and sampling adequacy.

Time Pressure and the Dilution Effect

Glover (1997) chose to examine the impact of time pressure and accountability using the dilution effect as a proxy for audit performance. He predicted that accountability would intensify the dilution effect; that accountable auditors would generate judgments of accounts receivable misstatement risk that are more moderate when compared to non-accountable auditors. In contrast to the accountability prediction, Glover (1997) also hypothesized that time

[8] When drive (stress) is increased, focus intensifies on cues relating to a central task compared to cues relating to a peripheral task (Easterbrook 1959).

pressure would decrease dilution, therefore improving audit performance. Results were found in support of the time pressure effect, but not the accountability effect. Therefore, auditors who were under time pressure were influenced less by nondiagnostic information, thus exhibiting less dilution; however accountability and the interaction between accountability and time pressure did not significantly impact dilution (Glover 1997).

The Glover (1994, 1997) findings can be explained by examining a series of psychology experiments on the impact of increased stress (electric shock) on performance (the learning of mice). Yerkes and Dodson (1908) found that the performance of mice on difficult tasks followed an inverted u-shaped curve. Learning improved as stress increased but only to a point, after which performance declined with increased stress. Broadhurst (1959), Anderson (1976), and Weick (1983) also describe the relationship of arousal (stress) and performance as an inverted u-shaped curve.

Easterbrook (1959) provides an additional explanation of the effect of stress on performance in the discussion of his cue-utilization theory. Easterbrook (1959) suggests that an increase in drive (stress) results in a smaller range of cues considered when forming a judgment. This could improve or harm performance. If greater stress reduces the use of irrelevant cues compared to the use of relevant cues, performance is improved (Easterbrook 1959). In the Glover (1997) study, findings suggest that time pressure (stress) is associated with a reduction in dilution, and an increase in auditor performance. For this to have occurred, the amount of stress must correspond to the portion of the inverted u-shaped curve that is increasing. Time pressure (stress) must not be beyond the stress midpoint on the Yerkes and Dodson (1908) curve, otherwise it would result in decreased performance. Therefore, in the

Glover (1994, 1997) study as time pressure increased, auditor performance improved. Glover (1997) surmised that this was due to auditors adopting a filtration strategy when there was less time available to complete the task. Increased time pressure enabled auditors to filter information that was not relevant to the judgment and allowed them to re-focus on the more important diagnostic client information.

Workload as an External Audit Pressure

Workload is another external pressure that can impact audit performance. Two archival studies (Lopez and Peters 2011; 2012) measure the association of workload pressures and audit outcomes, such as auditor switching and audit quality. Two separate variables are used to proxy for workload-related pressures, busy season workload and workload compression. Busy season workload is measured as an indicator variable with 1 assigned to clients with a fiscal year end of December, and 0 otherwise. Workload Compression, proxies for the external audit pressure due to fluctuations in the number of clients audited each month according to their fiscal year end. Audit fees are grouped by the fiscal year end month of the client. Fees for each month are summed and divided by the annual total of audit fees for that audit firm. An increase in this variable indicates increased auditor workload.

Evidence from Lopez and Peters (2011; 2012) has suggested that an increase in auditor workload pressures is associated with undesirable audit characteristics. For example, Lopez and Peters (2011) indicated that greater workload compression is associated with more auditor switching. In a subsequent study using the same methodology, Lopez and Peters (2012) found that when individual client characteristics are controlled, clients with a fiscal year-end month of

December are associated with lower audit quality, and as auditor workload compression increases, audit quality decreases.

These results seem to contradict the Glover (1997) finding that an increase in a different external pressure (time pressure) improved audit performance through the reduced influence of nondiagnostic client information on audit judgment. In contrast to Glover (1997), increased workload compression had a deleterious effect on audit quality. However, it is difficult to compare results across two very different methodologies, archival and experimental. Time pressure and workload may have opposite effects on auditor performance due to an unforeseen factor that may not have been controlled for in the archival model. One such factor might be auditor burnout.

Sweeney and Summers (2002) found that increased workload, proxied by the additional hours that public accountants[9] worked during busy season, was associated with increased burnout. In addition, when investigating the influence of burnout in public accounting, Fogarty et al. (2000) found evidence to support that burnout acts as a mediator between the role stressors and job outcomes of accountants. Fogarty et al. (2000) tested two different models, one with burnout and one without, but both containing the same "role stressors" ("role conflict," "role ambiguity," and "role overload") and job outcomes (job satisfaction, turnover intentions, and performance). No significant effect was found for role overload on job performance when burnout was not included in the model. However, tests of the model that contained burnout, exhibited evidence of a significant positive relationship between role overload and performance. However, when mediation tests were conducted, significant results

[9] Participants included public accountants of all experience levels in tax, audit and consulting (Sweeney and Summers 2002).

indicated that increased role overload increased burnout and burnout lowered job performance.

When considering the Fogarty et al. (2000) results, it is important to keep in mind that workload was not examined. Role overload is defined as, an unreasonably high degree of role requirements (Schick et al. 1990; Fogarty et al. 2000), and is not the same as increased workload. However, both role overload and increased workload place an unrealistic set of demands on the accountant or auditor. Therefore, the Fogarty et al. (2000) results raise the question of whether a similar relationship might exist when high workload and burnout exist in the same audit environment.

Decision Aids and Auditor Performance

Prior studies have provided evidence to suggest that decision aid use may have an impact on audit judgment. Wood (2012) found that the use of decision aids actually increased the influence of nondiagnostic information, therefore resulting in a greater dilution effect, and lower audit performance, when estimating the fraud risk of a client. These findings may be due to the type of decision aid that she used. The experimental task required that auditors in the decision aid condition answer a checklist of fraud risk factors. Wood (2012) suggested that this action guided them through the fraud risk judgment and resulted in auditors placing greater importance on dissimilar irrelevant client information compared to auditors with no decision aid.

Another study using a computerized decision aid examined the use of a checklist, logit model, and expert system on auditors management fraud risk estimates (Eining et al. 1997).

Their results suggested that auditor performance was the greatest for those that used the expert system decision aid. Additionally, those that used the logit model decision aid had greater performance than those that used the checklist and the unaided control group.

Hypotheses

Participants use a similarity based judgment when they compare a target's displayed characteristics to their internally held belief of the characteristics that a target should possess during a situation. For example, clients with characteristics that are similar to qualities displayed by low fraud-risk companies would be judged as having a lower risk of fraud (Hackenbrack 1992).

The use of similarity based judgments is complicated by the presence of nondiagnostic information (Hackenbrack 1992; Nisbett et al. 1981; Glover 1997). Nondiagnostic information do not directly relate to a judgment task, and therefore should not impact a participant's decision. However, prior studies, both in psychology and auditing, have shown that the addition of nondiagnostic information results in a significant change in a participant's judgment compared to when nondiagnostic information is absent (Nisbett et al. 1981; Glover 1997; Hackenbrack 1992; Hoffman and Patton 1997; Shelton 1999).

This study defines diagnostic client information as accounts receivable audit program results that directly influence an auditor's judgment of misstatement risk. Conversely, nondiagnostic client information is defined as information that is irrelevant to an assessment of accounts receivable misstatement risk, such as information about a client's inventory, payroll, or accounts payable (Glover 1994, 1997). H1 is designed to examine the impact of these two

different types of information on audit judgment. Support for this hypothesis will validate the results of prior nondiagnostic research (Nisbett et al. 1981; Glover 1997; Hackenbrack 1992; Hoffman and Patton 1997; Shelton 1999).

Dilution occurs when nondiagnostic information changes a participant's judgment according to a theorized pattern. Initial judgments, based only on diagnostic information, are more "extreme" (Nisbett et al. 1981, 256). When nondiagnostic information is introduced, initial judgments become diluted (Nisbett et al. 1981; Glover 1997).

> *H1: Misstatement risk judgments based on diagnostic and nondiagnostic information are significantly different from misstatement risk judgments based on diagnostic information only.*

Previous studies have shown that the amount of the dilution is responsive to an increase or decrease in multiple factors. For example, high time pressure conditions redirect auditor attention away from nondiagnostic information, thus resulting in a decrease in dilution (Glover 1994, 1997). However, when auditors are held accountable, their judgments of fraud-risk become more conservative, thus resulting in increased dilution (Hoffman and Patton 1997).

Time pressure has been studied extensively, but it is not the only job pressure that auditors encounter. An audit pressure that is routinely experienced in practice, but explored little in behavioral audit literature is workload. Results of archival studies suggest that an increase in auditor workload pressure is associated with undesirable audit characteristics, such as increased auditor switching and decreased audit quality (Lopez and Peters 2011; 2012).

Task load, beyond what can be reasonably completed in a specified time period, can also result in stress (McGrath 1976). Even if the individual tasks can be easily completed in

isolation, it is the increase in task load that results from the combination of multiple tasks that generates stress.

Therefore, unlike the relationship between time pressure and dilution, I propose that an increase in auditor workload (increase in the number of simultaneously audited clients) will increase dilution. In order to isolate the impact of workload, time pressure is not manipulated in this study. Instead, time pressure[10] is introduced to all participants at a constant rate.

> *H2: The dilution effect is greater when workload is high, than when workload is low.*[11]

According to theory, dilution occurs when initial judgments become "less extreme" in response to nondiagnostic information being added to a judgment task (Hackenbrack 1992, 126). The direction of the change in auditor judgment will depend on the auditor's initial judgment of the degree of misstatement risk; high risk versus low risk (Glover 1994; 1997). Initial judgments of high misstatement risk in high risk cases will dilute to secondary judgments of lower misstatement risk when nondiagnostic information is added. Conversely, low misstatement risk judgments in low risk cases will dilute to secondary judgments of higher misstatement risk when nondiagnostic information is added. This effect occurs because nondiagnostic information complicates decision making for auditors, causing their misstatement risk judgments to become diluted (Glover 1994; 1997). Support for this hypothesis will further validate the results seen in prior research.

[10] Time pressure is introduced at a moderately high level, slightly lower than that used in the Glover (1994, 1997) study. This has been chosen to simulate the level where Glover (1994, 1997) noted the greatest improvement on audit performance.

[11] H2 does not consider the direction of the change in auditor judgment when nondiagnostic information is added. Instead the absolute value of the mean dilution effect under high workload is compared with the absolute value of the mean dilution effect under low workload, regardless of whether the addition of nondiagnostic information resulted in a judgment that was more extreme or less extreme when compared to the initial judgment.

H3: The addition of nondiagnostic client information increases (decreases) misstatement risk judgments for low risk (high risk) audit cases.

Previous research in the area of audit performance and dilution has suggested that decision aids may be used to guide auditors away from undesired outcomes, such as the dilution of auditor judgments (Hoffman and Patton 1997, Hackenbrack 1992). However, Wood (2012) found evidence that suggests that decision aid use has the opposite effect. Wood (2012) hypothesized that irrelevant client information, not included in a decision aid checklist, heightened the auditor's perception of dissimilarity between the hypothetical client and a client depicting high fraud risk. Auditors interpreted this dissimilarity as an indication that the hypothetical client had a lower risk of fraud, which resulted in an increase in the dilution effect.

This study re-examines the impact of decision aids on dilution. Like Hoffman and Patton (1997) and Hackenbrack (1992), I predict that the use of a decision aid will reduce dilution.[12] Unlike Wood (2012), my PCAOB guidance "decision aid" does not involve a checklist, but is instead a reminder of the relevant audit standards related to accounts receivable misstatement risk. Auditors that review these audit standards prior to performing an audit task will have clear guidance of how accounts receivable should be audited, including which types of client information are relevant and which are irrelevant to the audit judgment. This guidance will enable auditors to compare the accounts receivable audit program results to client information contained in the audit guidance, therefore encouraging auditors to conduct a more accurate audit. As a result, auditors should focus their attention away from irrelevant client information and toward relevant client information, reducing dilution.

[12] Hoffman and Patton (1997) and Hackenbrack (1992) suggested that the use of a decision aid during fraud risk estimation might lessen dilution. This was suggested as an idea for future research in the discussion section.

H4: The dilution effect is reduced when PCAOB guidance is used, compared to when PCAOB guidance is not used.[13]

The following chapter discusses the methodology used for this study.

[13] H4 does not consider the direction of the change in auditor judgment when nondiagnostic information is added. Instead the absolute value of the mean dilution effect under PCAOB guidance is compared with the absolute value of the mean dilution effect under no PCAOB guidance, regardless of whether the addition of nondiagnostic information resulted in a judgment that was more extreme or less extreme when compared to the initial judgment.

CHAPTER 3

METHOLOGY

Variables

Independent variables include Client Information Type (diagnostic information, mixed information: diagnostic + nondiagnostic), Workload (low, high), and PCAOB Guidance (audit guidance, no audit guidance). The study uses a 2X2X2 mixed design, with time pressure held constant at a moderately high level, and workload and PCAOB guidance as the two manipulated independent variables. Client information type is a within subjects measure that is used to identify when dilution occurs.

Client Information Type varies based on whether the information is directly applicable to the audit misstatement risk assessment. Diagnostic client information is directly relevant to auditor judgment of client accounts receivable misstatement risk, while nondiagnostic client information is not directly relevant to auditor judgment. Therefore, Client Information Type is an indicator variable (1-diagnostic, 0-nondiagnostic).

According to the Glover (1994, 1997) findings, the dilution effect decreased when auditors experienced high time pressure, but did not disappear completely. Therefore in this study, time pressure is imposed across all conditions at a moderately high level using an efficiency preference statement, a technique similar to that used in Gramling (1999). It states,

> "*The audit partner has asked you to complete the audit work as quickly as possible, and has indicated that it should take you about 3 minutes to review the case and answer the following questions. However, you must continue to uphold the efficiency, profitability, and reputation motives of your firm as you perform the audits*."

Workload is manipulated at a high and low level to generate changes in the dilution effect. The workload independent variable is measured based on the number of clients included in Part 1 of the experimental task (see Figure 3.1), with low workload containing one client and high workload containing three clients (8 additional cases). Finally, the PCAOB Guidance independent variable is based on whether audit guidance is included in the experiment; audit guidance or no audit guidance.

Several studies in the audit performance literature have used dependent variables focused on the identification and collection of audit evidence (McDaniel 1990; Braun 2000; Bowrin 2001; Bowrin and King II 2010). Research in this area is important to identify factors that could have unwanted effects on audit evidence gathering. This study instead examines auditor performance in an earlier audit stage; the planning stage. The dilution effect of misstatement risk assessments made during the audit planning stage was chosen as the dependent variable for this study. Dilution draws strongly from key psychology literature such as Easterbrook's (1959) discussion of cue utilization, the Nisbett et al. (1981) psychology paper on dilution effect, and key psychology studies on the representativeness heuristic and similarity based judgments (Kahneman and Tversky 1972, 1973; Alba and Hutchinson 1987; Shelton 1999; Kahneman 2003; and Tversky 1977).

This study examines the dilution effect by comparing a participant's misstatement risk judgment on a diagnostic case with their misstatement risk judgment on the same case when

nondiagnostic client information is added. Dilution is tested within subjects.[14] However, the two independent variables, workload and PCAOB guidance, are varied between subjects.

Participants

Results of the Shelton (1999) dilution effect study indicated that inexperienced auditors (audit seniors) provided evidence of the dilution effect, while experienced auditors did not. Therefore, the two pilot questionnaires for this study used students as proxies for inexperienced audit staff. Students enrolled in financial accounting, tax, or audit, from a large university in the southwestern U.S. were surveyed. The questionnaires were administered to students in paper form. All four versions of the questionnaire were used in the second pilot study and randomization was achieved by alternating the four versions as the paper copies were handed out to students.

The final data collection utilized the online questionnaire site, Qualtrics. This allowed more complex randomizations to be built into the questionnaire flow[15] (see Figure 3.2) to reduce potential demand effects caused by the order that cases appear to participants. Participants (166) for the study were obtained from a variety of sources ranging from business and accounting alumni, to CPAs and auditors. Questionnaire data has been separated into three categories and analyzed to determine if significant differences in mean responses are associated between participant sources for the research sample.

[14] The dilution effect for high risk audit cases (Adgrow and KwikFoods) is measured as a participant's judgment on the diagnostic case minus the same participant's judgment on the mixed case, as seen in Glover (1994, 1997). The dilution effect for the low risk audit case (Medley) was calculated the same way. Absolute value of the dilution effect means are used for hypothesis testing.

[15] Participants from the alumni list randomly received one of the four versions of the questionnaire through an automatic process that was built into the questionnaire flow. Participants obtained through Qualtrics were randomly selected by Qualtrics to meet quotas for each of the four versions.

A small university in the southwestern U.S. was contacted and agreed to provide a personal email list of 623 accounting and business administration alumni from class years 1989 to the present. This list included accounting alumni of various professions, including CPAs and auditors (see demographics section). An invitation to take the questionnaire was emailed to these participants on April 9, 2015, which included a link to the Qualtrics questionnaire. After one week, a reminder email was sent out to all of the alumni on the list who had not yet submitted a questionnaire. One week later a total of 80 completed questionnaires had been submitted. Observations that were retained from this source were categorized as the Alumni Practitioner sample.

Invitations were then sent out to the work email of several of the more recent graduates from the small university with a request that they forward the questionnaire link to additional auditors that they might know. A recently hired faculty member at the same small university and a partner at a local audit firm were also sent similar requests, along with other auditor contacts available through the researcher's personal network. Additionally, an alumni member from a large university in the southwestern U.S. was contacted and the questionnaire link was posted on the university accounting alumni Facebook page. Observations that were collected as a result of these initiatives, and retained for data analysis, were assigned to a separate category for comparative analysis, called HSA (45).

Qualtrics was also used to obtain additional auditor responses, and a work order was signed requesting 70 completed questionnaires. Qualtrics contacted one of its panel partners who had identified auditor participants through a screening process. The questionnaire links were sent and eleven completed questionnaires were collected as a result of this soft launch.

Upon review of the data it was apparent that most of it could not be used. Several of the questionnaires were completed in less than eight minutes. Additionally, one participant indicated that he had assessed the risk of material misstatement for the net year-end accounts receivable balance 75,000 times. Qualtrics then conducted a second soft launch with a different panel partner. Seven questionnaires were collected from this attempt. Responses from six of the seven questionnaires were retained. However, one questionnaire response was discarded due to the participant being inattentive.

As a result of these difficulties, Qualtrics was contacted and the work order was canceled. A total of six questionnaire responses were collected. Observations retained from this source were added to the category named HSA.[16]

The researcher was also able to obtain a public list[17] of 6,904 licensed CPAs from a large metropolitan city in the southwestern U.S. An email invitation was sent to those on the list along with the questionnaire link[18]. Data collected from this source, and retained for data analysis, has been labeled with the category name, Practitioner (41).

Experimental Task/ Questionnaire

This study utilizes an accounts receivable audit task with a questionnaire instrument that is an adaptation of the questionnaire used in Glover's (1994) dissertation. The questionnaire contains three parts, along with demographic questions, and general questions

[16] HSA is a sample of participants who were referred to the questionnaire by participants from the other two samples; Practitioner and Alumni Practitioner. This sample category had 45 participants.
[17] This list was from the internet.
[18] See Chapter 4, Participant Demographics for information regarding number of responses and number retained in each sample category.

(see figure). The same diagnostic and mixed cases for all three clients (Adgrow, KwikFoods, and Medley) are included regardless of the experimental condition. The questionnaires vary on level of workload (low, high) and PCAOB reminder (absent, present), creating four different versions of the questionnaire instrument (high workload x PCAOB; high workload x no PCAOB; low workload x PCAOB; low workload x no PCAOB). The workload manipulation does not change the content of the questionnaire, instead it changes the order in which the cases are presented to the audit participant (see figure 3.1). Adgrow and KwikFoods cases include diagnostic information that is indicative of a client with a high risk of misstatement. On the contrary, Medley cases depict diagnostic client information that is low risk. Both high and low risk cases are included in this study to verify that dilution is occurring in both scenarios.

Two types of audit cases are included in the questionnaire, diagnostic cases and mixed cases. Diagnostic cases are one page in length and include an accounts receivable audit program containing four audit procedures along with short descriptions of the audit results to date. On the same page, below the accounts receivable audit program, are three questions that a participant must answer. These one-page cases are diagnostic because all of the client information contained within the case is directly relevant to the accounts receivable audit. Mixed cases are also included in the questionnaire and contain a combination of diagnostic and nondiagnostic client information along with the same three questions that are included in the diagnostic cases. The diagnostic information is an exact replica of one of the one-page diagnostic cases already included in the questionnaire. The nondiagnostic information is client information that is not relevant to the accounts receivable audit. The three mixed cases used in the study (mixed cases for Adgrow, KwikFoods, and Medley) were created by splitting the

Glover (1994) mixed case into three shorter mixed cases. All nondiagnostic audit programs have been taken from Glover (1994) and are identical in all three mixed cases. The nondiagnostic information pretested and validated by Glover (1994) has been used in order to strengthen the internal validity of the study.

Workload (low, high) is operationalized as the number of clients audited in Part 1 of the questionnaire instrument. The low workload condition only contains one client (Adgrow), while the high workload condition contains three clients (Adgrow, KwikFoods, and Medley). Each client that is present in Part 1 of the questionnaire contains four one-page diagnostic cases.

Part 1 of the low workload condition includes four of the eight Adgrow, Inc. diagnostic[19] cases that Glover (1994) used. The four that were selected for Part 1 of this study are identical to the four Glover (1994) diagnostic cases that are of higher accounts receivable misstatement risk. These four are randomly ordered and contain the same four audit procedures which vary according to the audit results. After reviewing the accounts receivable audit program including the 2013 accounts receivable audit evidence to date, auditors are first asked to provide a rating in response to the statement, "given the evidence available at this point in the audit, assess the risk that the net year-end accounts receivable balance is materially misstated" on a 9-point Likert scale from minimum risk to maximum risk. This same question was included in the Glover (1994) study and is the question that will be used to examine the dilution effect.

[19] Cases are diagnostic when they contain only client information that is relevant to the audit judgment being made (Glover 1994, 1997).

Secondly, auditors are asked to respond to, "the ***audit procedures*** contained in the ***accounts receivable audit program*** are sufficient to assess the risk that the net year-end accounts receivable balance is materially misstated" on a 9-point Likert scale from strongly disagree to strongly agree. Finally, auditors must also respond to the statement, "how confident are you of the accuracy of your risk assessment?" on a 9-point Likert scale with points corresponding to 1-not confident, 5-somewhat confident, and 9-very confident; also used in Glover (1994). Audit participants answer the same three questions for each of the four one-page Adgrow diagnostic cases and each of these cases are considered independently of each other.

Part 1 of the high workload condition includes the same four Adgrow, Inc. diagnostic cases used in the low workload condition plus four KwikFoods diagnostic cases and four Medley diagnostic cases. The four KwikFoods diagnostic cases are similar to four of the Adgrow diagnostic cases used in Glover (1994); only the audit results have been slightly changed. The four Medley diagnostic cases are identical to the four diagnostic Adgrow cases with the lower accounts receivable misstatement risk that were used in Glover (1994). Each of the twelve one-page diagnostic cases have the same three 9-point Likert scale questions discussed above.

Demographic questions are included in a separate section between Part 1 and Part 2 of the questionnaire instrument (see Figure 3.1). Participants are asked to provide answers regarding their education, professional experience, and demographic characteristics. Data analysis has been conducted to determine if significant differences in response can be attributed to differences in audit experience.

Part 2 contains one Adgrow, Inc. mixed case (diagnostic and nondiagnostic client information) that was adapted from the mixed case used in Glover (1994). It contains 2013 year end client information including client summary and background information, summary of operations, inventory, payroll, subsequent events, an accounts payable audit program, and a property, equipment, and depreciation audit program. All of this information is considered nondiagnostic because it is not relevant to the accounts receivable misstatement risk judgment that participating auditors are making. The nondiagnostic information contained in the Adgrow mixed case (as well as the KwikFoods and Medley mixed cases) has been adapted from Glover (1994). Additionally, the mixed case includes one of the Glover (1994) Adgrow diagnostic cases used in Part 1 of the study corresponding to the case with the highest accounts receivable misstatement risk. There is no difference in the content of the Adgrow, Inc. mixed cases across the low workload and high workload conditions, however a one-page excerpt from the PCAOB audit standards was only included in the PCAOB-present conditions.

Part 3 of the low workload condition includes cases for both KwikFoods and Medley. There are four one-page KwikFoods diagnostic cases followed by a KwikFoods mixed case. Then there are four one-page Medley diagnostic cases followed by a Medley mixed case. Each of these cases are structured in the same manner as those described above. Part 3 of the high workload condition only includes the KwikFoods mixed case and the Medley mixed case. Auditors in this condition have already answered the diagnostic cases for all three clients in Part 1 of the questionnaire. The KwikFoods and Medley mixed cases included in Part 3 are the same regardless of the workload condition. However, an additional page of PCAOB audit standards is included in the PCAOB – present conditions only.

There is a short section of manipulation check questions following Part 3 of the questionnaire. All participants must answer the same questions regardless of the experimental condition.

Auditor judgments based on diagnostic information only will be compared to auditor judgments based on diagnostic and nondiagnostic information, to confirm that nondiagnostic information results in a dilution effect (Glover 1997). Then the impact of the two manipulated factors (workload and PCAOB guidance) will be analyzed to determine their effect on dilution.

Figure 3.1

Questionnaire Instrument Design

PCAOB

Low Workload		High Workload	
Part 1	Adgrow Diagnostic Case 1 (AdDC1Q1) Adgrow Diagnostic Case 2 Adgrow Diagnostic Case 3 Adgrow Diagnostic Case 4	Part 1	Adgrow Diagnostic Case 1 (AdDC1Q1) Adgrow Diagnostic Case 2 Adgrow Diagnostic Case 3 Adgrow Diagnostic Case 4 KwikFoods Diagnostic Case 1 (KwDC1Q1) KwikFoods Diagnostic Case 2 KwikFoods Diagnostic Case 3 KwikFoods Diagnostic Case 4 Medley Diagnostic Case 1 (MdDC1Q1) Medley Diagnostic Case 2 Medley Diagnostic Case 3 Medley Diagnostic Case 4
Demographic	Demographic Questions	Demographic	Demographic Questions
Part 2	Adgrow Mixed Case (AdMCQ1) **PCAOB Guidance Inside Mixed Case**	Part 2	Adgrow Mixed Case (AdMCQ1) **PCAOB Guidance Inside Mixed Case**
Part 3	KwikFoods Diagnostic Case 1 (KwDC1Q1) KwikFoods Diagnostic Case 2 KwikFoods Diagnostic Case 3 KwikFoods Diagnostic Case 4 KwikFoods Mixed Case (KwMCQ1) **PCAOB Guidance Inside Mixed Case** Medley Diagnostic Case 1 (MdDC1Q1) Medley Diagnostic Case 2 Medley Diagnostic Case 3 Medley Diagnostic Case 4 Medley Mixed Case (MdMCQ1) **PCAOB Guidance Inside Mixed Case**	Part 3	KwikFoods Mixed Case (KwMCQ1) **PCAOB Guidance Inside Mixed Case** Medley Mixed Case (MdMCQ1) **PCAOB Guidance Inside Mixed Case**
General Ques	Manipulation Check Questions	General Ques	Manipulation Check Questions

No PCAOB

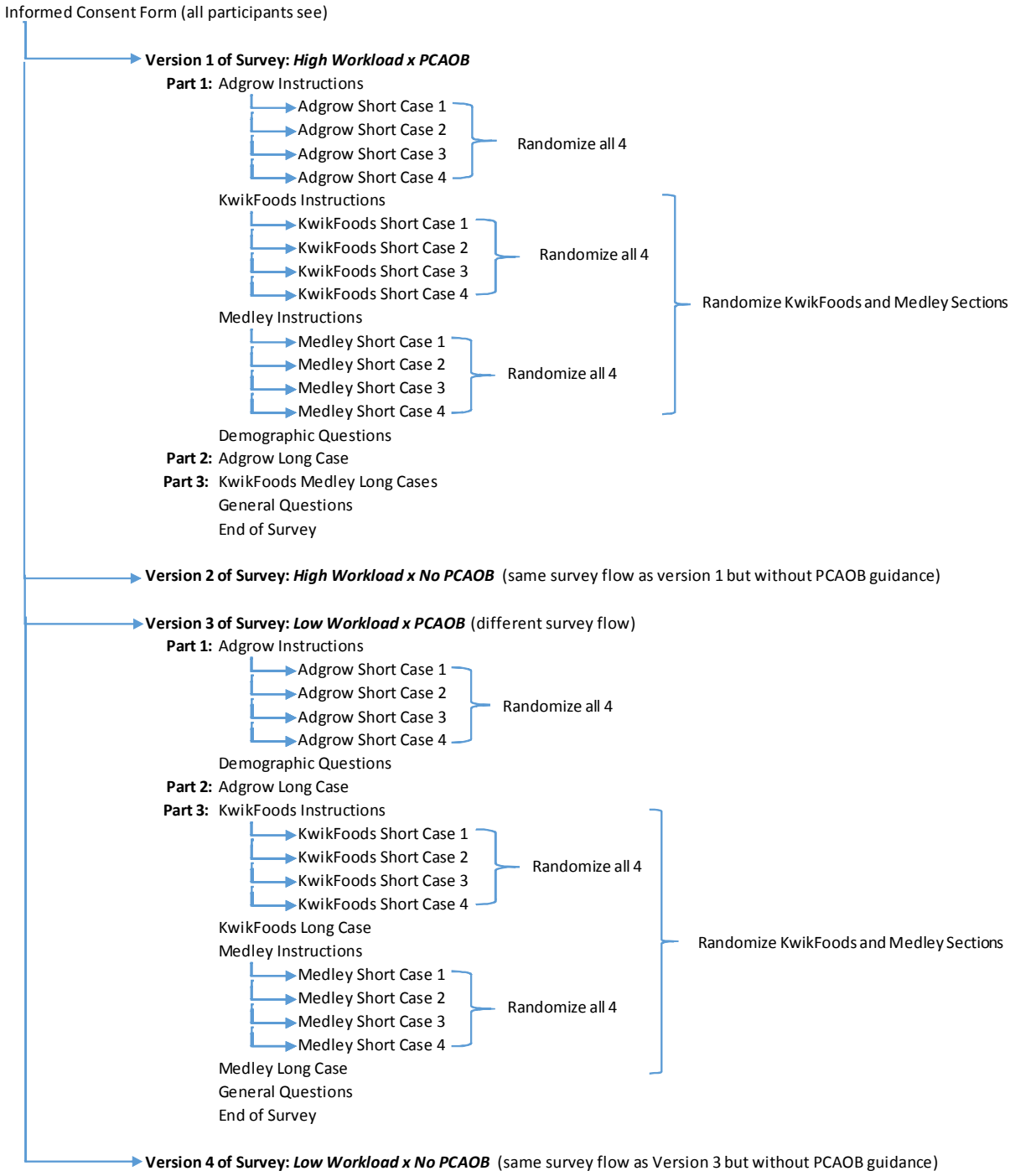
Low Workload		High Workload	
Part 1	Adgrow Diagnostic Case 1 (AdDC1Q1) Adgrow Diagnostic Case 2 Adgrow Diagnostic Case 3 Adgrow Diagnostic Case 4	Part 1	Adgrow Diagnostic Case 1 (AdDC1Q1) Adgrow Diagnostic Case 2 Adgrow Diagnostic Case 3 Adgrow Diagnostic Case 4 KwikFoods Diagnostic Case 1 (KwDC1Q1) KwikFoods Diagnostic Case 2 KwikFoods Diagnostic Case 3 KwikFoods Diagnostic Case 4 Medley Diagnostic Case 1 (MdDC1Q1) Medley Diagnostic Case 2 Medley Diagnostic Case 3 Medley Diagnostic Case 4
Demographic	Demographic Questions	Demographic	Demographic Questions
Part 2	Adgrow Mixed Case (AdMCQ1) No PCAOB Guidance	Part 2	Adgrow Mixed Case (AdMCQ1) No PCAOB Guidance
Part 3	KwikFoods Diagnostic Case 1 (KwDC1Q1) KwikFoods Diagnostic Case 2 KwikFoods Diagnostic Case 3 KwikFoods Diagnostic Case 4 KwikFoods Mixed Case (KwMCQ1) No PCAOB Guidance Medley Diagnostic Case 1 (MdDC1Q1) Medley Diagnostic Case 2 Medley Diagnostic Case 3 Medley Diagnostic Case 4 Medley Mixed Case (MdMCQ1) No PCAOB Guidance	Part 3	KwikFoods Mixed Case (KwMCQ1) No PCAOB Guidance Medley Mixed Case (MdMCQ1) No PCAOB Guidance
General Ques	Manipulation Check Questions	General Ques	Manipulation Check Questions

* Adgrow and KwikFoods Diagnostic cases 1-4 have a high risk of material misstatement

** Medley Diagnostic cases 1-4 have a low risk of material misstatement

Figure 3.2

Questionnaire Randomization: Used in Final Collection



CHAPTER 4

RESULTS

Descriptive Statistics

This section contains a detailed discussion of the demographic characteristics of the data, including means, standard deviations, and frequencies.

Participant Demographics

Questionnaire data was collected during the period April 9, 2015 to May 26, 2015 using Qualtrics Questionnaire Software. A total of 166 observations were collected from three different sources; Practitioner (41), HSA (45), and Alumni Practitioner (80). Nine observations were dropped because one or more variables that were necessary to calculate the dilution effect were missing. This resulted in a total of 157 usable observations.

Participants were asked a series of questions pertaining to the highest level of education completed, experience, and general demographic characteristics. With regard to their education, the majority of participants had a CPA license (120 or 76.92% of participants; see Table 4.1), and had earned a master's degree (97 or 61.78% of participants; see Table 4.2).

Table 4.1

CPA License

Level	Frequency	Percent
Licensed CPA	120	76.92
Not Licensed CPA	36	23.08

Frequency Missing = 1

Table 4.2

Highest Level of Education Completed

Level	Frequency	Percent
Bachelor's Degree	44	28.03
Some Graduate Work	9	5.73
Master's Degree	97	61.78
JD	5	3.18
PhD	1	0.64
Other	1	0.64

Their work experience was more varied (see Table 4.3). The job role of 38 participants (24.2%) was that of an external auditor, and 6 (3.82%) were internal auditors. 49 participants (31.41%) are in the audit department at their firm. Males made up a larger percentage of participants at 59.24% (93 participants) compared to females at 40.76% (64 participants – see Table 4.4).

Table 4.3

Job Role

Level	Frequency	Percent
CPA	22	14.01
Accounting Clerk	3	1.91
External Auditor	38	24.2
Internal Auditor	6	3.82
Tax Accountant	11	7.01
Controller	14	8.92
CFO	10	6.37
Other	53	33.76

Table 4.4

Gender

Level	Frequency	Percent
Male	93	59.24
Female	64	40.76

Mean Responses

The mean age of participants was 38.65 years. On average they had spent 7.77 years at their present job, and 5.90 years in auditing. They had audited accounts receivable about 20.87 times and had assessed the risk of material misstatement for the net year-end accounts receivable balance about 12.86 times (see Table 4.5). See Appendix A for other variable means.

Table 4.5

Demographic Variables
Means and Standard Deviations

Variable	N	Mean	Standard Deviation
Yr(s) in Job	155	7.7652	8.6769
Yr(s) in Auditing	156	5.9015	8.5440
Time(s) Audited Manuf. Client	147	4.1769	13.5998
Time(s) Audited Food Processing Client	152	0.6316	2.8533
Time(s) Audited Accounts Receivable	138	20.8696	47.0707
Time(s) Assessed Risk of Material Miss.	139	12.8561	29.4841
Age	156	38.6474	11.8691

Differences in Sample Demographics

Three separate samples were collected in order to obtain sufficient participants for this study. A list of accounting student alumni from a small university in the southwestern U.S. were invited via email to participate in the study and a questionnaire link was provided in the email.

Out of 623 students contacted, 80 submitted the questionnaire which yielded a response rate of 80/623 = 12.84%. Four questionnaires were incomplete and were therefore dropped from analysis resulting in an Alumni Practitioner sample of 76 completed questionnaires. A list of accounting professionals from a large metropolitan city in the southwestern U.S. also received an email request for participation. Out of 6,904 professionals contacted, 41 questionnaires were submitted from this sample (response rate of 41/6,904 = 0.59%). After dropping three questionnaires due to incomplete data, 38 were retained. Additionally, 45 questionnaires were collected as a result of the Alumni Practitioner and the Practitioner participants forwarding the invitation email to their colleagues. Two of these questionnaires were incomplete and therefore dropped, resulting in a HSA sample size of 43 completed questionnaires.

In order to determine if observations from these three sample sources can be combined for analysis, demographic variables were tested to determine if significant differences exist between samples. The observations were coded according to their sample source (1=Practitioner, 2=Alumni Practitioner, 3=HSA), then chi-square testing was performed on the categorical demographic data to test for significant differences. Results of chi-square testing indicate that job role and gender are not significantly different (at the 5% alpha level) between the three sample groups, as indicated by p-values of 0.0719 and 0.2449 respectively (see Table 4.6). Significant differences between sample groups do not exist for auditor title (chi-square p-value of 0.1597). However, results indicate that this variable had 33% of cells with expected counts of less than 5. This violates an important assumption of chi-square testing. In order for it to be an appropriate method of analysis the data must not have any expected frequencies of less than one and expected frequencies of less than five should not exceed 20% (Michael 2001).

If this assumption is violated, as in the chi-square test of auditor title, it is more appropriate to perform a Fisher's Exact Test. Therefore, a Fisher's Exact Test was conducted, resulting in a p-value of 0.1687, still indicating that significant differences in auditor title do not exist across sample groups.

Table 4.6

Demographics by Sample

	Practitioner		Alumni Practitioner		H.S.A.		Chi-square	Fisher's Exact
	N	mean	N	mean	N	mean	p-value	p-value
Education	38	3.0526	76	3.6447	43	3.4419	0.0440	0.0377
Job Role	38	4.8947	76	5.5000	43	4.5349	0.0719	n/a
Department	38	3.4211	76	4.0395	42	2.9762	0.0028	n/a
CPA	38	1.0000	75	1.3333	43	1.2558	0.0003	n/a
Auditor Title	38	3.7105	75	4.1467	41	3.6829	0.1597	0.1687
Firm Size	38	4.2632	75	4.9467	43	4.1860	0.0071	n/a
Gender	38	1.3684	76	1.4737	43	1.3256	0.2449	n/a

Additionally, ANOVA testing was conducted to test for significant differences in continuous demographic variables due to the sample source. Results indicated that significant differences between sample groups do not exist for the number of times the participant assessed the risk of material misstatement for net year-end accounts receivable balance, as indicated by an F-value of 2.68 and a p-value of 0.0721 (see Table 4.7). Because the Practitioner sample, Alumni Practitioner sample, and the HSA sample are not significantly different for job role, auditor title, gender, and the number of times the participant assessed the risk of material misstatement for net year-end accounts receivable balance observations from these three samples can be combined into one sample for hypothesis testing.

Table 4.7

Demographics by Sample

	Practitioner		Alumni Practitioner		H.S.A.		ANOVA	
	N	mean	N	mean	N	mean	F Value	p-value
Yr in JobVar	38	9.7697	76	5.8993	43	9.0698	3.38	0.0367
Yr in AuditVar	38	8.0658	76	3.8153	42	7.9564	4.92	0.0085
Num A/RVar	35	40.2286	69	11.2899	35	21.2286	4.64	0.0112
Num Miss RiskVar	34	21.1471	68	7.5735	38	15.8684	2.68	0.0721
Age	38	46.9211	76	34.2763	42	39.0714	17.48	<.0001

Nonresponse Testing

To test for nonresponse bias, chi-square tests were conducted on demographic indicator variables, and t-tests were run on continuous demographic variables. Because the data for each sample was collected at different times, not concurrently, nonresponse testing was performed for each sample separately. This was done to test for demographic differences between the first two thirds of respondents and the last third of respondents (Armstrong and Overton 1977). Before chi-square tests were run, the demographic variables were transformed, combining certain levels in order to minimize the incidence of cells with expected counts of less than 5. This was done to reduce the chance that a chi-square test assumption would be violated due to small expected frequencies.

Chi-square tests on the Practitioner sample indicated that a significant difference was found for Department and Firm Size, with p-values of 0.0146 and 0.0110 respectively (at the alpha=.05 level; see Table 4.8). However, SAS 9.3 produced a note stating that both of the variables had “63% of the cells with expected counts of less than 5,” potentially rendering the chi-square tests invalid. As a result, a Fisher's Exact Test was also run for each; yielding p-values of 0.0115 and 0.0094 respectively (see Table 4.14). The results of these tests indicate a

potential for non-response bias for department and firm size. This may have occurred because some departments were busier during the time of the questionnaire. This may have resulted in fewer participants responding from those firms, or taking longer to respond.

Table 4.8

Nonresponse Bias Test
Chi-squares Using Transformed Variables

	Practitioner0: first 2/3			Practitioner1: last 1/3			chi-square
Variable	**N**	**Mean**	**Std Dev**	**N**	**Mean**	**Std Dev**	**p-value**
Educ	25	3.1600	0.9866	13	2.8462	0.9871	**0.5729**
CPA	25	1.0000	0.0000	13	1.0000	0.0000	**n/a**
JobRol	25	4.2800	2.5904	13	6.0769	1.7541	**0.0951**
Dept	25	2.8400	1.7243	13	4.5385	1.6641	**0.0146**
AudTitl	25	3.5600	1.4457	13	4.0000	1.4142	**0.3325**
FirmSiz	25	3.6000	2.2361	13	5.5385	1.3914	**0.0110**
Gender	25	1.3600	0.4899	13	1.3846	0.5064	**0.8814**

T-Tests for the Practitioner sample did not result in any p-values that were lower than .05, but instead yielded p-values ranging from 0.3267 to 0.8851 (see Table 4.9). These results indicate that the number of years in the present job, years in auditing, times auditing accounts receivable, times assessed the risk of material misstatement for net year-end accounts receivable balance, and the age of participants are not significantly different between early and late respondents.

Table 4.9

Nonresponse Bias Test

	Practitioner0: first 2/3			Practitioner1: last 1/3			t-test	t-test
Variable	**N**	**Mean**	**Std Dev**	**N**	**Mean**	**Std Dev**	**t-value**	**p-value**
YrinJobVar	25	10.3700	12.1906	13	8.6154	6.6494	0.57	**0.5696**
YrinAudVar	25	8.2600	12.4391	13	7.6923	8.9943	0.15	**0.8851**
NumARVar	24	46.0833	80.1221	11	27.4545	30.1176	1.00	**0.3267**
NumMissVar	23	21.9565	39.6169	11	19.4545	31.2197	0.18	**0.8556**
Age	25	46.0800	15.1957	13	48.5385	11.8435	-0.51	**0.6149**

Testing of the Alumni Practitioner sample resulted in chi-square p-values of 0.0423 for Auditor Title, and 0.0126 for Firm Size (see Table 4.10). Fisher's Exact Tests were also run for these variables due to cells with expected counts of less than 5 (70% and 50% respectively, see Table 4.14). The results of the Fisher's Exact Tests still showed a potential for non-response bias with p-values of 0.0432 for Auditor Title and 0.0130 for Firm Size. This could be due to a delay in questionnaire response for auditors with a higher audit title.

Table 4.10

Nonresponse Bias Test
Chi-squares Using Transformed Variables

	Alumni Practitioner0: first 2/3			Alumni Practitioner1: last 1/3			chi-square
Variable	**N**	**Mean**	**Std Dev**	**N**	**Mean**	**Std Dev**	**p-value**
Educ	50	3.7600	0.7969	26	3.4231	1.0648	**0.2169**
CPA	49	1.3265	0.4738	26	1.3462	0.4852	**0.8638**
JobRol	50	5.3600	2.7237	26	5.7692	2.5504	**0.3181**
Dept	50	4.1000	2.0628	26	3.9231	2.0770	**0.9395**
AudTitl	50	4.1600	1.4337	25	4.1200	1.2014	**0.0423**
FirmSiz	50	5.2200	1.5687	25	4.4000	1.9579	**0.0126**
Gender	50	1.4200	0.4986	26	1.5769	0.5038	**0.1937**

The Alumni Practitioner sample also did not show significant differences in the continuous demographic variables due to response timing, as indicated by t-test results with p-values higher than .05 (see Table 4.11).

Table 4.11

Nonresponse Bias Test

	Alumni Practitioner0: first 2/3			Alumni Practitioner1: last 1/3			t-test	t-test
Variable	**N**	**Mean**	**Std Dev**	**N**	**Mean**	**Std Dev**	**t-value**	**p-value**
YrinJobVar	50	5.6070	5.0682	26	6.4615	6.5068	-0.63	**0.5296**
YrinAudVar	50	3.4112	3.3509	26	4.5923	5.0686	-1.07	**0.2905**
NumARVar	47	10.8085	21.2354	22	12.3182	21.1916	-0.28	**0.7839**
NumMissVar	48	7.6042	20.6930	20	7.5000	14.6161	0.02	**0.9837**
Age	50	33.9600	7.8088	26	34.8846	8.7422	-0.47	**0.6397**

Chi-square tests and Fisher's Exact Tests for the HSA sample did not result in p-values less than .05 (see Tables 4.12 and 4.14). Therefore, it can be assumed that the demographic variables in this sample do not have a potential for non-response bias. However, t-tests suggested that nonresponse bias might exist in a participants number of years in auditing (t-value = 2.58, p-value = 0.0139, see Table 4.13).

Table 4.12

Nonresponse Bias Test
Chi-squares Using Transformed Variables

	HSA0: first 2/3			HSA1: last 1/3			chi-square
Variable	**N**	**Mean**	**Std Dev**	**N**	**Mean**	**Std Dev**	**p-value**
Educ	29	3.4138	1.1807	14	3.5000	0.8549	**0.7179**
CPA	29	1.2069	0.4123	14	1.3571	0.4972	**0.29**
JobRol	29	4.6897	2.8922	14	4.2143	1.9287	**0.4387**
Dept	28	3.0714	2.1244	14	2.7857	2.2250	**0.7558**
AudTitl	27	3.7778	1.2195	14	3.5000	1.7431	**0.2151**
FirmSiz	29	3.8966	1.9883	14	4.7857	1.3688	**0.0706**
Gender	29	1.3448	0.4837	14	1.2857	0.4688	**0.6983**

Table 4.13

Nonresponse Bias Test

	HSA0: first 2/3			HSA1: last 1/3			t-test	t-test
Variable	**N**	**Mean**	**Std Dev**	**N**	**Mean**	**Std Dev**	**t-value**	**p-value**
YrinJobVar	29	10.7759	11.5369	14	5.5357	7.4848	1.54	**0.1301**
YrinAudVar	28	10.2143	12.2213	14	3.4407	4.6942	2.58	**0.0139**
NumARVar	23	26.1739	63.6279	12	11.7500	20.6359	0.99	**0.3294**
NumMissVar	24	20.6250	44.0067	14	7.7143	13.1350	1.34	**0.1910**
Age	28	40.2500	12.1461	14	36.7143	10.6874	0.92	**0.3611**

Table 4.14

Nonresponse Bias Test
Fisher's Exact Using Transformed Variables

	Practitioner	Alumni Practitioner	H.S.A.
	p-value	p-value	p-value
Education	**0.4618**	**0.1843**	**0.8816**
CPA	n/a	n/a	**0.4568**
Job Role	**0.0950**	n/a	**0.4692**
Department	**0.0115**	n/a	**0.7564**
Auditor Title	**0.3347**	**0.0432**	**0.2197**
Firm Size	**0.0094**	**0.0130**	**0.0908**
Gender	**1.0000**	n/a	**1.0000**

Test for Differences between Accountant Types

A one-way MANOVA[20] was conducted to test for differences in auditor responses across accountant types. An indicator variable for auditor job role was created and coded a one if the participant was an auditor (external or internal), or zero otherwise. Results suggest that there are no significant differences in auditor response between auditors and non-auditors, as indicated by the following: F (18, 135) = 1.236, p = .242; Wilks' Lambda = .859; partial eta squared = .141. Pillai's Trace was also examined due to unequal sample sizes. All values were the same, except Pillai's Trace = .141. Results of the one-way MANOVA indicate that non-auditors can be retained in the sample.

Prior to hypothesis testing three participants were dropped from the data set. These participants identified themselves as accounting clerks. The auditor task in this research study requires professional accounting and/or auditor experience. Therefore, the clerks would not be qualified to participate in this study.

[20] Three clerks were dropped and a reduced sample of 154 participants were used for the one-way MANOVA test.

Results

Hypothesis Testing

Hypothesis[21] 1 states that misstatement risk judgments based on diagnostic and nondiagnostic information are significantly different from misstatement risk judgments based on diagnostic information only. Finding support for this hypothesis is one step toward validating the findings of prior research such as Glover (1994, 1997). A dependent group t-test was used to pair participants' answers on diagnostic case one with their answer on the corresponding mixed case. P-values of less than .05 on question one from Adgrow and Medley (Adgrow: p-value = <.0001; Medley: p-value = <.0001) indicate that a significant difference exists in a participant's misstatement risk assessment between the two cases (see Table 4.15). Because the only difference between the two cases is the addition of nondiagnostic information in the mixed case, these question one results provide support for H1 for Adgrow and Medley.

Hypothesis 2 proposes that the dilution effect is greater when workload is high than when workload is low.[22] An independent group t-test was conducted to test for a significant difference in the dilution effect variable (AdDilEffQ1ABS, KwDilEffQ1ABS, and MdDilEffQ1V2ABS) across high and low workload conditions for all three clients. Because the dilution effect variable is a measure of the change in an auditor's judgment the absolute value of the dilution effect data was used in the analysis for H2 and H4, and the related supplemental analyses. Significant differences were found for the KwikFoods and Medley

[21] After the demographic analyses in the preceding section were conducted, five participants that were sourced through Qualtrics were dropped. All hypothesis and supplemental hypothesis testing were conducted with a sample of 149 participants.

[22] H2 and H4 do not consider the direction of the change in auditor judgment when nondiagnostic information is added. Instead the absolute value of the mean dilution effect is used for H2 and H4 testing.

dilution effect question one variable (see Table 4.16). The dilution effect means for KwikFoods and Medley follow the pattern proposed in H2. Under high workload conditions the mean absolute value of the dilution effect was 1.1707, and 1.3293 for KwikFoods and Medley clients respectively, when compared to 0.8209, and 0.8657 for the same clients under low workload conditions.

Table 4.15

H1 & H3: Dependent Group T-test with All Data (paired)

		N	DC1 mean	Std Dev	N	MC mean	Std Dev	t-value	two-tailed p-value	one-tailed p-value
Adgrow (high risk)	**Question 1**	149	6.2349	1.4630	149	5.6913	1.4973	4.62	<.0001	<.00005
	Question 2	149	5.5570	2.0807	149	5.7584	1.8911	-1.80	0.0739	0.0370
	Question 3	149	6.0403	1.6761	149	6.0067	1.6002	0.37	0.7084	0.3542
KwikFoods (high risk)	**Question 1**	149	6.0940	1.7059	149	5.8725	1.4251	1.89	0.0611	0.0306
	Question 2	149	5.7785	1.9131	149	5.9329	1.8183	-1.51	0.1340	0.0670
	Question 3	149	6.2013	1.5853	149	6.0336	1.6250	2.06	0.0415	0.0208
Medley (low risk)	**Question 1**	149	2.0671	1.4269	149	2.9597	1.8595	-6.89	<.0001	<.00005
	Question 2	149	6.5705	2.1631	149	6.2953	1.9539	2.00	0.0469	0.0235
	Question 3	149	7.3020	1.6673	149	6.7248	1.6310	5.16	<.0001	<.00005

Table 4.16

H2: Independent Group T-test with All Data (Workload)

		High Workload Dilution Effect (ABS)			Low Workload Dilution Effect (ABS)			t-value	one-tailed p-value
		N	mean	Std Dev	N	mean	Std Dev		
Adgrow	**Question 1**	82	1.0732	1.0157	67	1.1493	1.1181	0.43	0.3322
KwikFoods	**Question 1**	82	1.1707	1.1526	67	0.8209	0.8335	-2.15	0.0168
Medley	**Question 1**	82	1.3293	1.6029	67	0.8657	1.1401	-2.06	0.0207

Hypothesis 3 states that the addition of nondiagnostic client information increases (decreases) misstatement risk judgments for low risk (high risk) audit cases (see Table 4.15). Dependent group t-test results show a significant difference in a participant's risk assessment

between the diagnostic case and its corresponding mixed case, (Adgrow: p-value = <.00005; KwikFoods: p-value = 0.0306; and Medley: p-value = <.00005, one-tailed).[23] The directional nature of the movement between the two participant's answers supports the dilution effect theory and validates the results of prior research; the mean of the second risk judgment is lower for high risk cases (Adgrow and KwikFoods) and higher for the low risk case (Medley).

Hypothesis 4 proposes that the dilution effect is reduced when PCAOB guidance is used, compared to when it is absent. An independent group t-test was conducted to test for a significant difference in the mean dilution effect (AdDilEffQ1ABS, KwDilEffQ1ABS, and MdDilEffQ1V2ABS) between participants that did not review PCAOB guidance and those who did. No significant differences were found at the alpha 5% level (see Table 4.17). Even though significance was absent, there was evidence of the proposed effect of PCAOB guidance on the mean absolute value of the dilution effect for KwikFoods and Medley. Under the no PCAOB condition, the mean absolute value of the dilution effect was 1.1375 and 1.1750 for KwikFoods and Medley respectively compared to 0.8696 and 1.0580, when PCAOB guidance was used.

Table 4.17

H4: Independent Group T-test with All Data (PCAOB)

		PCAOB Dilution Effect (ABS)			No PCAOB Dilution Effect (ABS)			t-value	one-tailed p-value
		N	mean	Std Dev	N	mean	Std Dev		
Adgrow	**Question 1**	69	1.1884	1.0884	80	1.0375	1.0366	-0.87	0.1940
KwikFoods	**Question 1**	69	0.8696	0.9535	80	1.1375	1.0878	1.59	0.0574
Medley	**Question 1**	69	1.0580	1.5135	80	1.1750	1.3574	0.5	0.3098

[23] These results are from the same analysis used to test H1. P-values from Table 4.15 have been halved because H3 was a directional hypothesis; one-tailed.

Supplemental Analyses

All Adgrow and KwikFoods cases possess a high risk of accounts receivable material misstatement. Supplemental analyses were conducted to re-test all four hypotheses using the average of the Adgrow and KwikFoods response variables (AdKwik). Question one measures the auditor's assessment of the risk that the year-end accounts receivable balance is materially misstated. For the dilution effect to occur there should be a significant difference between the auditor's first response to question one and their second response to question one when, nondiagnostic information is added (H1). Table 4.18 shows support for this hypothesis with a p-value of <.0001. Furthermore, for high risk cases, dilution effect theory predicts that the auditor's response will move from a higher risk of accounts receivable material misstatement to a lower risk of accounts receivable misstatement when nondiagnostic information is added (H3). Comparing the means in Table 4.18 show support for this hypothesis with the mean accounts receivable misstatement risk moving from 6.1644 to 5.7819 (p-value of <.00005,[24] one-tailed).

Table 4.18

H1 & H3: Dependent Group T-test with Adgrow and KwikFoods Averaged (paired)

		DC1			MC				two-tailed	one-tailed
		N	mean	Std Dev	N	mean	Std Dev	t-value	p-value	p-value
AdKwik (high risk)	Question 1	149	6.16443	1.433719	149	5.7819	1.319452	4.25	<.0001	<.00005
	Question 2	149	5.66779	1.87138	149	5.8456	1.73585	-2.1	0.0376	0.0188
	Question 3	149	6.12081	1.524187	149	6.0201	1.517777	1.64	0.1039	0.05195

[24] H3 is a directional hypothesis. The p-value from Table 4.18 H1 was halved because H3 is one-tailed.

H2 was also re-tested using the absolute value of the average[25] of the mean dilution effect for Adgrow and KwikFoods. A significant difference in dilution effect was found across workload conditions (high and low) for question two (p-value = 0.0314). The trend in means for all three questions follow theory, the absolute value of the mean dilution effect is greater under high workload conditions than under low workload conditions (see Table 4.19).

Table 4.19

H2: Independent Group T-test Adgrow and KwikFoods Averaged (Workload)

		High Workload Dilution Effect (ABS)			Low Workload Dilution Effect (ABS)			t-value	one-tailed p-value
		N	mean	Std Dev	N	mean	Std Dev		
AdKwik (high risk)	Question 1	82	0.9268	0.8503	67	0.7612	0.7037	-1.28	0.1019
	Question 2	82	0.7561	0.9726	67	0.5149	0.5772	-1.88	0.0314
	Question 3	82	0.5122	0.6284	67	0.4478	0.5232	-0.67	0.2518

H4 tests were also conducted with the absolute value of the averaged Adgrow and KwikFoods response variables. No significant difference was found, as indicated by p-values that exceeded .05 (see Table 4.20). However, a trend in means for question one and three follows theory.

Table 4.20

H4: Independent Group T-test Adgrow and KwikFoods Averaged (PCAOB)

		PCAOB Dilution Effect (ABS)			No PCAOB Dilution Effect (ABS)			t-value	one-tailed p-value
		N	mean	Std Dev	N	mean	Std Dev		
AdKwik (high risk)	Question 1	69	0.8261	0.7565	80	0.8750	0.8210	0.38	0.3538
	Question 2	69	0.7536	0.9800	80	0.5563	0.6560	-1.42	0.0791
	Question 3	69	0.4565	0.5126	80	0.5063	0.6389	0.52	0.3025

[25] Before supplemental tests were conducted for H2 and H4, the Adgrow dilution effect and Kwikfoods dilution effect was averaged for each question (question 1, 2, and 3 separately). The absolute values of the new averaged variables were calculated and then used in statistical testing.

Discussion

Results of hypothesis testing for H1 and H3 validate prior studies of the impact of nondiagnostic information on participant judgments. When nondiagnostic information was added auditor judgments changed, and the two were significantly different. Validation of this effect was seen for auditor responses to question one for Adgrow and Medley. Results of H3 testing, for question one, confirm the theoretical foundations of dilution effect theory for all three clients; that judgments that are initially high will be reduced and those that are initially low will be increased when nondiagnostic information is added. This nondiagnostic information effect can be explained by a variety of theories.

One possible explanation is a participants' use of similarity based judgments, a comparison of how similar a target is to its desired outcome (Kahneman and Tversky 1972; 1973). Tversky (1977) suggested that a target that includes mostly relevant information will be viewed as more similar to the desired outcome, or internally held belief, while those with mostly irrelevant information will be viewed as dissimilar therefore resulting in a re-evaluation of the participant's judgment. In this study, relevant information, referred to as diagnostic information, are the results of the accounts receivable audit program included in each case. These results have a direct impact on an audit participants' judgment of accounts receivable misstatement risk. Judgments made using this information alone result in an initial evaluation of accounts receivable misstatement risk.

Irrelevant information, referred to as nondiagnostic information, includes other client information, which, although important to other areas of the audit, is not important to the audit judgment task in this study. When auditors' judge accounts receivable misstatement risk

based on a combination of diagnostic and nondiagnostic information, the presence of the nondiagnostic information makes the case appear different. Therefore, the auditor is prompted to change their initial risk judgment.

The results of H3 testing pertaining to the direction of the auditors' change in risk judgment can be explained by Nisbett et al. (1981). Results of this psychology study suggested that including diagnostic information only, resulted in a perception that the target was more similar to an internally held stereotype. This resulted in judgments that were relatively more "extreme" when compared to judgments based on nondiagnostic and diagnostic information (Nisbett et al. 1981, 256). The cases in this study that included diagnostic information only (results from the accounts receivable audit program) were perceived by auditors as more similar to their internal stereotype of a high risk or low risk client. This perception resulted in initial risk judgments that were higher for high risk cases, and lower for low risk cases, and secondary judgments that were diluted. The direction of the secondary judgments followed the theorized direction, a reduction of high risk judgments and an increase of low risk judgments. Therefore, it was the nondiagnostic information that resulted in a change of the auditors' risk judgment, as a reaction to a perceived reduction in the similarity between the subject and an internally held stereotype.

Regarding H2 and H4, changes in the absolute value of the dilution effect means for several clients trended in the proposed direction (although significance was not always present). These changes can be attributed to the impact of stress or pressure on performance. One of the earliest theories describing this relationship comes from a fundamental psychology study by Yerkes and Dodson (1908). The results of this study suggested that as a stimulus in

increases, learning improves. However, this effect only occurs up to a certain level of stimulus. Once it is increased beyond that level, learning is hampered. Subsequent research in this area began to refer to this relationship as the inverted u-shaped curve between arousal (stress) and performance (Broadhurst 1959; Anderson 1976; and Weick 1983). Accounting research in the 1980s and 1990s extended this theory to examine the relationship between time pressure and auditor judgment. Glover (1994, 1997) found evidence to support that an increase in time pressure can improve audit performance by lessening the effect of nondiagnostic information on auditor judgment. This study extends the investigation further by analyzing the effect of workload pressure on auditor judgment.

The results of H2 testing show trends indicating that pressure can have deleterious effects on performance as well. Assimilating the contrasting results of this study with Glover (1994, 1997) lends credence to the theory of the inverted u-shaped curve between pressure and performance. Increased pressure can result in either an improvement of performance or a reduction in performance depending on the degree of pressure. It can be theorized that Glover's level of time pressure was not great enough to push over the apex of the curve, while the level of pressure evoked by the increase in workload was.

An alternate explanation for the results of this study can be explained by the Easterbrook (1959) cue-utilization theory. Easterbrook (1959) suggested that increased stress results in the use of a smaller range of cues when forming a judgment. He explained that a reduction in the use of irrelevant cues compared to the use of relevant cues, improves performance. Glover (1994, 1997) saw this as a potential explanation for the results of his study. However, the opposite effect can also occur. If increased stress results in more irrelevant

cues being considered, relative to relevant cues, performance will suffer. This could also provide an important perspective on this study's results. Increased workload could have resulted in auditors paying closer attention to nondiagnostic information, relative to diagnostic information, thus diluting their judgment of misstatement risk.

Limitations

Several limitations to this study exist. First, access to professional auditors was limited. Although many participants identified themselves as external or internal auditors (44/157 = 28.03%),[26] the majority were not practicing auditors. This may have resulted in participants who lacked relevant auditing experience. Although the results of Shelton (1999) suggest that only inexperienced auditors exhibit dilution, participants must have significant knowledge of audit-related and misstatement risk issues in order to effectively complete the audit judgment tasks in this study. Participants that had never worked as a professional auditor,[27] or had not for many years, would more than likely skew the data, therefore causing a lack of support for the proposed hypotheses.

Secondly, due to low response rates, several samples of participants were used. All three samples were gathered within about a month, therefore reducing potential differences in response due to longitudinal time effects. However some differences in demographic data were detected across sample groups. Therefore, it would strengthen future studies to obtain all participants from the same sample group.

[26] This is the ratio of auditors to total calculated during demographic testing, when the total sample equaled 157.
[27] Most of the participants in this study had at least worked as an audit intern during college.

Finally, the small sample size may be a substantial limitation to this study and may have contributed to the lack of statistical significance. Testing of the H2 and H4 hypotheses resulted in a noticeable reduction of the sample size per condition. As a result enough power might not have been present to test for significance.

Contributions

This study contributes both to audit judgment literature, as well as audit practice. Statistical support for H1 and H3 validate the findings of prior audit dilution effect research, while clearly providing evidence in support of the directional nature of the dilution effect theory. Although statistical support is not consistently found for the effects of workload pressure on auditor judgments, analyses provide evidence of the proposed trending (trending is also seen for PCAOB). The effects of nondiagnostic client information and the trends in auditor judgment seen across workload and PCAOB conditions may help auditors be more cognizant of how easily auditor judgment can be influenced.

Appendix A

Cell 1

Means and Standard Deviations for Low Workload x No GAAS Observations
Adgrow, Inc.

Variable	N[28]	Mean	Standard Deviation
AdDC1Q1	35	6.2857	1.3842
AdDC1Q2	35	5.3714	2.1974
AdDC1Q3	35	5.9429	1.6259
AdDC2Q1	35	5.9143	1.4627
AdDC2Q2	35	5.6857	1.8435
AdDC2Q3	35	5.8571	1.7002
AdDC3Q1	35	5.2286	1.6285
AdDC3Q2	35	5.8857	1.7619
AdDC3Q3	35	6.0571	1.5135
AdDC4Q1	35	4.3714	1.8325
AdDC4Q2	35	6.0571	1.8620
AdDC4Q3	35	6.3714	1.3080
AdMCQ1	35	5.6286	1.6104
AdMCQ2	35	5.6	1.9584
AdMCQ3	35	5.9429	1.2113
AdDilEffQ1	35	0.6571	1.3708
AdDilEffQ2	35	-0.229	1.3738
AdDilEffQ3	35	0	1.0572

Cell 5

Means and Standard Deviations for Low Workload x No GAAS Observations
KwikFoods, Inc.

Variable	N	Mean	Standard Deviation
KwDC1Q1	35	5.7714	1.6465
KwDC1Q2	35	5.8857	1.7111
KwDC1Q3	35	6.0286	1.3391
KwDC2Q1	35	5.4000	1.4593
KwDC2Q2	35	5.9429	1.6793
KwDC2Q3	35	6.2571	1.2210
KwDC3Q1	35	4.3714	1.8164
KwDC3Q2	34	5.8235	1.7662
KwDC3Q3	35	6.3714	1.5163
KwDC4Q1	35	4.8286	1.7403
KwDC4Q2	35	5.8857	1.7619
KwDC4Q3	35	6.2286	1.1398
KwMCQ1	35	5.6571	1.4741
KwMCQ2	35	5.8286	1.6357
KwMCQ3	35	5.6571	1.5328
KwDilEffQ1	35	0.1143	1.4302
KwDilEffQ2	35	0.0571	1.2353
KwDilEffQ3	35	0.3714	0.9727

Cell 9

Means and Standard Deviations for Low Workload x No GAAS Observations
Medley, Inc.

Variable	N	Mean	Standard Deviation
MdDC1Q1	35	2.2857	2.0230
MdDC1Q2	35	6.9143	1.7213
MdDC1Q3	35	7.1714	1.3609
MdDC2Q1	35	4.6571	1.5519
MdDC2Q2	35	6.1143	1.5862
MdDC2Q3	35	6.3714	1.3080
MdDC3Q1	35	3.6000	1.7690
MdDC3Q2	35	6.3143	1.4506
MdDC3Q3	35	6.6286	1.2387
MdDC4Q1	35	4.7429	1.3793
MdDC4Q2	35	5.9714	1.4448
MdDC4Q3	35	6.1429	1.2866
MdMCQ1	35	3.0571	1.9545
MdMCQ2	35	6.5714	1.7704
MdMCQ3	35	6.6286	1.5920
MdDilEffQ1V2	35	-0.7714	1.0870
MdDilEffQ2V2	35	0.3429	1.3048
MdDilEffQ3V2	35	0.5429	1.1718

[28] Descriptive statistics (in Appendix A) were conducted on the 157 participant sample.

Cell 2

Means and Standard Deviations for Low Workload x Yes GAAS Observations

Adgrow, Inc.

Variable	N	Mean	Standard Deviation
AdDC1Q1	35	6.1714	1.4849
AdDC1Q2	35	6.0857	1.9307
AdDC1Q3	35	6.3429	1.6968
AdDC2Q1	35	5.4857	1.7884
AdDC2Q2	35	6	2.2096
AdDC2Q3	35	6.2857	1.6009
AdDC3Q1	35	4.4857	1.8048
AdDC3Q2	35	6.2571	1.8840
AdDC3Q3	35	6.0286	1.7738
AdDC4Q1	35	4.7429	2.1191
AdDC4Q2	35	6.1143	2.0972
AdDC4Q3	35	6.5143	1.7042
AdMCQ1	35	5.6857	1.5862
AdMCQ2	35	6.0286	1.9476
AdMCQ3	35	6.2286	1.7837
AdDilEffQ1	35	0.4857	1.6156
AdDilEffQ2	35	0.0571	1.3048
AdDilEffQ3	35	0.1143	1.0508

Cell 6

Means and Standard Deviations for Low Workload x Yes GAAS Observations

KwikFoods, Inc.

Variable	N	Mean	Standard Deviation
KwDC1Q1	35	6.0286	1.8066
KwDC1Q2	35	6.0857	1.9307
KwDC1Q3	35	6.5429	1.8043
KwDC2Q1	35	5.6000	1.7523
KwDC2Q2	35	6.1429	1.9272
KwDC2Q3	35	6.4857	1.6337
KwDC3Q1	35	4.5429	1.7714
KwDC3Q2	34	6.2647	1.8473
KwDC3Q3	35	6.5714	1.5584
KwDC4Q1	35	4.7714	1.8485
KwDC4Q2	34	6.2941	1.8179
KwDC4Q3	35	6.4857	1.6693
KwMCQ1	35	5.8000	1.5107
KwMCQ2	35	6.1714	1.9018
KwMCQ3	35	6.2000	1.7790
KwDilEffQ1	35	0.2286	1.0314
KwDilEffQ2	35	-0.0857	0.7811
KwDilEffQ3	35	0.3429	0.8023

Cell 10

Means and Standard Deviations for Low Workload x Yes GAAS Observations

Medley, Inc.

Variable	N	Mean	Standard Deviation
MdDC1Q1	35	2.4	1.8818
MdDC1Q2	35	6.8857	1.8907
MdDC1Q3	35	7.5714	1.5957
MdDC2Q1	35	4.9143	1.6337
MdDC2Q2	35	6.1143	2.0113
MdDC2Q3	35	6.4000	1.6485
MdDC3Q1	35	4.0857	1.8845
MdDC3Q2	35	6.4286	1.8034
MdDC3Q3	35	6.5429	1.6863
MdDC4Q1	35	4.9714	1.9325
MdDC4Q2	35	6.3143	1.8274
MdDC4Q3	35	6.2857	1.5825
MdMCQ1	35	2.9714	2.0071
MdMCQ2	35	6.3714	2.1016
MdMCQ3	35	7.0571	1.7647
MdDilEffQ1V2	35	-0.5714	1.3781
MdDilEffQ2V2	35	0.5143	2.2012
MdDilEffQ3V2	35	0.5143	1.0675

Cell 3

Means and Standard Deviations for High Workload x No GAAS Observations

Adgrow, Inc.

Variable	N	Mean	Standard Deviation
AdDC1Q1	50	6.1400	1.7025
AdDC1Q2	50	5.5200	2.2337
AdDC1Q3	50	5.8400	1.7654
AdDC2Q1	50	5.3200	1.6218
AdDC2Q2	50	5.7000	2.2246
AdDC2Q3	50	6.0800	1.5887
AdDC3Q1	50	4.5000	1.7173
AdDC3Q2	50	5.5800	2.3655
AdDC3Q3	50	6.2400	1.6971
AdDC4Q1	50	4.2000	1.4569
AdDC4Q2	50	5.3800	2.2305
AdDC4Q3	50	5.9600	1.8178
AdMCQ1	50	5.7800	1.4609
AdMCQ2	50	5.7800	1.9197
AdMCQ3	50	5.8800	1.8142
AdDilEffQ1	50	0.3600	1.3667
AdDilEffQ2	50	-0.2600	1.5361
AdDilEffQ3	50	-0.0400	1.1945

Cell 7

Means and Standard Deviations for High Workload x No GAAS Observations

KwikFoods, Inc.

Variable	N	Mean	Standard Deviation
KwDC1Q1	50	6.08	1.8165
KwDC1Q2	50	5.66	2.0958
KwDC1Q3	50	6.1	1.7409
KwDC2Q1	50	5.54	1.7168
KwDC2Q2	50	5.7	2.0429
KwDC2Q3	50	6.16	1.6704
KwDC3Q1	50	4.74	1.6759
KwDC3Q2	50	5.82	1.9659
KwDC3Q3	50	6.22	1.6073
KwDC4Q1	50	4.54	1.7404
KwDC4Q2	50	5.6	2.0996
KwDC4Q3	50	6.16	1.5434
KwMCQ1	50	5.86	1.4144
KwMCQ2	50	5.74	1.8495
KwMCQ3	50	6	1.6288
KwDilEffQ1	50	0.22	1.7296
KwDilEffQ2	50	-0.08	1.1400
KwDilEffQ3	50	0.1	1.0926

Cell 11

Means and Standard Deviations for High Workload x No GAAS Observations

Medley, Inc.

Variable	N	Mean	Standard Deviation
MdDC1Q1	50	2.0000	1.3851
MdDC1Q2	50	6.4400	2.3918
MdDC1Q3	50	7.0800	2.0086
MdDC2Q1	50	4.3000	1.5152
MdDC2Q2	50	5.6600	2.0860
MdDC2Q3	50	6.1000	1.6933
MdDC3Q1	50	3.3400	1.3940
MdDC3Q2	50	5.7200	2.2043
MdDC3Q3	49	6.4286	1.7795
MdDC4Q1	50	4.6400	1.6383
MdDC4Q2	50	5.8000	2.0702
MdDC4Q3	49	6.1224	1.6025
MdMCQ1	50	3.0400	1.9162
MdMCQ2	50	6.1000	1.9509
MdMCQ3	50	6.5800	1.6424
MdDilEffQ1V2	50	-1.0400	1.8622
MdDilEffQ2V2	50	0.3400	1.4930
MdDilEffQ3V2	50	0.5000	1.6067

Cell 4

Means and Standard Deviations for High Workload x Yes GAAS Observations

Adgrow, Inc.

Variable	N	Mean	Standard Deviation
AdDC1Q1	37	6.59459	1.2124232
AdDC1Q2	37	5.24324	1.8915808
AdDC1Q3	37	6.2973	1.5431725
AdDC2Q1	37	5.62162	1.5697019
AdDC2Q2	37	5.24324	1.9494359
AdDC2Q3	37	5.81081	1.7925438
AdDC3Q1	37	5	1.7950549
AdDC3Q2	37	5.40541	2.1661469
AdDC3Q3	37	6.10811	1.5235883
AdDC4Q1	37	4.97297	1.5181584
AdDC4Q2	37	5.02703	2.0342414
AdDC4Q3	37	5.97297	1.6747551
AdMCQ1	37	5.83784	1.5002502
AdMCQ2	37	5.72973	1.6939213
AdMCQ3	37	6.21622	1.3971227
AdDilEffQ1	37	0.75676	1.3417527
AdDilEffQ2	37	-0.48649	1.3868754
AdDilEffQ3	37	0.08108	0.982581

Cell 8

Means and Standard Deviations for High Workload x Yes GAAS Observations

KwikFoods, Inc.

Variable	N	Mean	Standard Deviation
KwDC1Q1	37	6.7297297	1.3672488
KwDC1Q2	37	5.6486486	1.8441532
KwDC1Q3	37	6.4054054	1.279311
KwDC2Q1	37	5.7297297	1.6098421
KwDC2Q2	37	5.5135135	1.9238506
KwDC2Q3	37	6.4324324	1.4051599
KwDC3Q1	37	5.0810811	1.6223847
KwDC3Q2	37	5.6216216	1.8760383
KwDC3Q3	37	6.2162162	1.3151919
KwDC4Q1	37	5.0540541	1.8400777
KwDC4Q2	37	5.6486486	1.8888006
KwDC4Q3	37	6.4054054	1.4617239
KwMCQ1	37	6.2702703	1.3874166
KwMCQ2	37	6.0540541	1.8995021
KwMCQ3	37	6.3513514	1.4947355
KwDilEffQ1	37	0.4594595	1.4258447
KwDilEffQ2	37	-0.4054054	1.9359101
KwDilEffQ3	37	0.0540541	1.1041826

Cell 12

Means and Standard Deviations for High Workload x Yes GAAS Observations

Medley, Inc.

Variable	N	Mean	Standard Deviation
MdDC1Q1	37	2.3243243	1.6508256
MdDC1Q2	37	6.1891892	2.2710477
MdDC1Q3	37	7.5135135	1.325427
MdDC2Q1	37	4.972973	1.6070416
MdDC2Q2	37	5.4594595	2.0762492
MdDC2Q3	37	6.027027	1.4811123
MdDC3Q1	37	4.0540541	1.7150631
MdDC3Q2	37	5.6486486	2.0168212
MdDC3Q3	37	6.5135135	1.4836446
MdDC4Q1	37	4.7297297	1.4841505
MdDC4Q2	37	5.6756757	1.8567303
MdDC4Q3	37	6.027027	1.423737
MdMCQ1	37	3.3783784	2.2029192
MdMCQ2	37	6.3513514	1.8888006
MdMCQ3	37	6.8378378	1.4242642
MdDilEffQ1V2	37	-1.0540541	1.8096339
MdDilEffQ2V2	37	-0.1621622	1.5186528
MdDilEffQ3V2	37	0.6756757	1.4917189

REFERENCES

Alba, J. W., and J. W. Hutchinson. 1987. Dimensions of consumer expertise. *Journal of Consumer Research* 13: 411-454.

American Institute of Certified Public Accountants (AICPA). 2012. *Planning an Audit.* Statement on Auditing Standards No. 122, AU-C Section 300, .04. New York, NY: AICPA.

Anderson, C. R. 1976. Coping Behaviors as Intervening Mechanisms in the Inverted-U Stress-Performance Relationship. *Journal of Applied Psychology*: 30-34.

Armstrong and Overton 1977. Estimating Nonresponse Bias in Mail Surveys. *Journal of Marketing Research* XIV (August): 396-402.

Bowrin, A. R. 2001. "Time Pressure, Task Structure, and Audit Effectiveness." Dissertation.

Bowrin, A. R., and J. King II. 2010. Time pressure, task complexity, and audit effectiveness. *Managerial Auditing Journal* 25 (2): 160-181.

Braun, R. L. 1994. "The Effect of Time Pressure on Auditor Attention to Red Flags in a Dual-task Environment." Dissertation.

Braun, R. L. 2000. The effect of time pressure on auditor attention to qualitative aspects of misstatements indicative of potential fraudulent financial reporting. *Accounting, Organizations & Society* 25 (3): 243-259.

Broadhurst, P. L. 1959. The Interaction of Task Difficulty and Motivation: The Yerkes-Dodson Law Revived. *Acta Psychologica* 16: 321-338.

Choo, Freddie. 1995. Auditors' judgment performance under stress: A test of the predicted relationship by three theoretical models. *Journal of Accounting, Auditing & Finance* 10 (3): 611-641.

DeZoort, F. T., and A. T. Lord. 1997. A Review and Synthesis of Pressure Effects Research in Accounting. *Journal of Accounting Literature* 16: 28-85.

Easterbrook, J. A. 1953. The measurement of perseverance. Unpublished M. A. thesis, Queen's University Library, Kingston, Ontario

Easterbrook, J. A. 1959. The effect of emotion on cue utilization and the organization of behavior. *Psychological Review* 66 (3): 183-201.

Eining, M. M., D. R. Jones, and J. K. Loebbecke. 1997. Reliance on Decision Aids: An Examination of Auditors' Assessment of Management Fraud. *Auditing: A Journal of Practice and Theory* 16: 1-19.

Fogarty, T. J., Singh, J., Rhoads, G. K., and Ronald K. Moore. 2000. Antecedents and Consequences of Burnout in Accounting: Beyond the Role Stress Model. *Behavioral Research In Accounting* 12: 31-67.

Glover, S. M. 1994 "The Influence of Accountability and Time Pressure on Auditor Judgment." Dissertation.

Glover, S. M. 1997. The Influence of Time Pressure and Accountability on Auditors' Processing of Nondiagnostic Information. *Journal of Accounting Research* 35 (2): 213-226.

Gramling, A. A. 1999. External auditors' reliance on work performed by internal auditors: The influence of fee pressure on this reliance decision. *Auditing* 18: 117.

Hackenbrack, K. 1992. Implications of Seemingly Irrelevant Evidence in Audit Judgment. *Journal of Accounting Research* 30 (1): 126-136.

Hoffman, V. B., and J. M. Patton. 1997. Accountability, the Dilution Effect, and Conservatism in Auditors' Fraud Judgments. *Journal of Accounting Research* 35 (2): 227-237.

International Auditing and Assurance Standards Board (IAASB). 2011. *Enhancing the value of auditor reporting: Exploring options for change.* New York, NY: International Federation of Accountants (IFAC).

Kahneman, D., and A. Tversky 1972. Subjective probability: A judgment of representativeness. *Cognitive Psychology* 3: 430-454.

Kahneman, D., and A. Tversky 1973. On the psychology of prediction. *Psychological Review* 80: 237-251.

Kahneman, D. 2003. Maps of Bounded Rationality: Psychology for Behavioral Economics. *American Economic Review* 93 (5): 1449-1475.

López, D. M., and G. F. Peters. 2011. Auditor Workload Compression and Busy Season Auditor Switching. *Accounting Horizons* 25 (2): 357-380.

López, D. M., and G. F. Peters. 2012. The Effect of Workload Compression on Audit Quality. *Auditing: A Journal of Practice and Theory* 31 (4): 139-165.

Low, K., and H. Tan. 2011. Does Time Constraint Lead to Poorer Audit Performance? Effects of Forewarning of Impending Time Constraints and Instructions. *Auditing: A Journal of Practice & Theory* 30 (4): 173-190.

Margheim, L., T. Kelley, and D. Pattison. 2005. An Empirical Analysis of the Effects of Auditor Time Budget Pressure and Time Deadline Pressure. *Journal of Applied Business Research* 21 (1): 23-35.

McDaniel, L. S. 1990. The Effects of Time Pressure and Audit Program Structure on Audit Performance. *Journal of Accounting Research* 28 (2): 267-285.

McGrath, J. E. 1976. "Stress and Behavior in Organizations." *Handbook of Industrial and Organizational Psychology* 1351-1396.

Michael, Robert S. "Crosstabulation & chi square." *Indiana University, Bloomington, IN. URL http://www. indiana. edu/~ educy520/sec5982/we ek_12/chi_sq_summary011020. pdf (Visited 2010, June 15)* (2001).

Nisbett, R. E., H. Zukier, and R. E. Lemley. 1981. The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology* 13 (2): 248-277.

Otley, D. and A. Fakiolas 2000. Reliance on Accounting Performance Measures: Dead or a new beginning? *Accounting, Organizations and Society* 25 (4): 497-510.

Schick, A., Gordon, L., and S. Haka. 1990. Information Overload: A temporal approach. *Accounting, Organizations and Society* 15 (2): 199-220.

Shelton, S. W. 1994. "The Effect of Nondiagnostic Evidence and Accountability on Auditor Judgment." Dissertation.

Shelton, S. W. 1999. The Effect of Experience on the Use of Irrelevant Evidence in Auditor Judgment. *The Accounting Review* 74 (2): 217-224.

Sweeney, J. T., and S. L. Summers. 2002. The Effect of the Busy Season Workload on Public Accountants' Job Burnout. *Behavioral Research In Accounting* 14: 223-245.

Tversky, A. 1977. Features of similarity. *Psychological review* 84: 327-352.

Weick, K. E. 1983. Stress in Accounting Systems. The Accounting Review 58 (2): 350-369.

Wood, L. I. 2012. The impact of decision aid use on the dilution effect when assessing fraud. *Journal of Finance and Accountancy* 9: 23-42.

Yerkes, R.M., and J.D. Dodson. 1908. The Relation of Strength of Stimulus to Rapidity of Habit Formation. *Journal of Comparative Neurology & Psychology* 18: 459-482.