Why Do Auditors Over-Rely on Weak Analytical Procedures? The Role of Outcome Bias and Insensitivity to Precision

Steven M. Glover
Glover@byu.edu
(801) 422-6080

Douglas F. Prawitt
Doug_Prawitt@byu.edu
(801) 422-2351

T. Jeffrey Wilks
Jeff_Wilks@byu.edu
(801) 422-1728

School of Accountancy and Information Systems
Marriott School of Management
Brigham Young University
Provo, Utah 84602

March 11, 2004

Why Do Auditors Over-Rely on Weak Analytical Procedures? The Role of Outcome Bias and Insensitivity to Precision

Abstract

Recent evidence from highly publicized frauds and from the Public Oversight Board's Panel on Audit Effectiveness (2000) indicates that auditors sometimes over-rely on weak evidence supporting a clean audit opinion. In experiment 1 we examine whether a favorable outcome bias, together with insensitivity to factors relating to the strength of evidence, could lead to over-reliance on weak substantive procedures. In experiment 2 we examine whether an explicit *ex ante* cue can sensitize auditors to the weaknesses of an unreliable substantive procedure. Our examination is conducted in the context of substantive analytical procedures that are performed to provide evidence during the substantive testing phase of an audit.

The first experiment provides evidence that auditors attribute more strength to a weak, aggregate-level analytical procedure that produces an expectation that is not significantly different from the unaudited numbers (i.e., a "*favorable"* outcome), than to the same analytical procedure that produces an expectation that is significantly different from the unaudited numbers. Further, auditors in the "favorable outcome" condition adjust their assessments down to a level very similar to that of auditors in the "unfavorable outcome" condition after receiving further information about the relevant underlying state of nature, while auditors in the "unfavorable outcome" condition do not change from their initial assessment after receiving the additional information. Our results suggest that when an low-quality analytical procedure yields no significant difference, auditors tend to overestimate the strength of the evidence provided as compared to assessments when the same procedure yields a significant difference, and as compared to subsequent re-assessments of the same procedure after receiving additional relevant information. The second experiment provides evidence that auditors cued to explicitly evaluate the weaknesses in an aggregate procedure prior to calculating the results attribute less strength to that procedure than auditors who are not cued *ex ante*. However, graduate auditing students who are similarly cued attribute the same amount of strength to the aggregate procedure as students who are not cued, suggesting that the intervention depends on auditors' knowledge of the determinants of precision.

Key Terms: Outcome bias, evidence quality, substantive analytical procedures, evidence assessment, over-reliance, auditor judgment, financial statement analysis.

Data Availability: Contact the authors.

Why Do Auditors Over-Rely on Weak Analytical Procedures? The Role of Outcome Bias and Insensitivity to Precision

I. INTRODUCTION

In December 2003, the press first reported the story of the massive Parmalat, Inc. accounting fraud. As part of the fraud, the company claimed to have $4.9 billion on deposit at the Bank of America, but the money did not exist. While it is standard practice for auditors to confirm balances directly with the bank, in Parmalat's case, the auditors apparently chose to rely on a faxed copy of a fraudulent $4.9 billion confirmation fabricated by Parmalat accountants. The facts available at this point suggest that the auditors inappropriately relied on weak substantive evidence. This example of over-reliance on weak audit evidence is not an isolated instance. In 2000, the Public Oversight Board's (POB) Panel on Audit Effectiveness reported that in 20 percent of the situations where analytical procedures were used as the primary substantive test in a given area, the analytical procedures were too weak to provide the level of assurance required under the circumstances.

Why do auditors sometimes over-rely on weak and unreliable audit evidence? While there are likely multiple contributing factors, in this study we examine whether a favorable outcome bias, together with insensitivity to factors relating to the strength of evidence, could lead to over-reliance on weak substantive procedures. We also examine whether an explicit *ex ante* cue can sensitize auditors to the weaknesses of an unreliable substantive procedure. Our examination is conducted in the context of substantive analytical procedures that are performed to provide evidence during the substantive testing phase of an audit.

The purpose of substantive analytical procedures is to obtain assurance, sometimes in combination with other substantive testing, that accounts are fairly stated (Arens and

Loebbecke 2000, 189). Although prior research shows that auditors' judgments of evidence reliability are generally sensitive to the competence and objectivity of the evidence source (Bamber 1983; Hirst 1994; Caster and Pincus 1996), auditors' judgments regarding the quality of the evidence provided by substantive analytical procedures have received little attention. Given that auditors frequently rely on analytical procedures as the primary substantive test in an audit area (Ameen and Strawser 1994; Hirst and Koonce 1996), and given the POB's (2000) finding that auditors too often over-rely on weak analytical procedures in these situations, it is important to understand the factors that influence auditors' perceptions of the strength of the evidence provided by substantive analytical procedures.

In this study, we conduct two experiments focusing on a setting in which a material misstatement exists, but in which a weak analytical procedure either does or does not provide results indicating the presence of the material misstatement. In this setting, inappropriate reliance on weak analytical procedures suggesting no material misstatement may compromise audit effectiveness. In the first experiment, we examine whether a failure to consider the expectation precision[1] of an analytical procedure in conjunction with an outcome-based bias—the tendency to over-estimate the strength of the evidence when a "*favorable"* outcome is observed—can explain, in part, why auditors sometimes inappropriately rely on weak analytical procedures as documented by the POB. In the second experiment we also examine whether, in the presence of a favorable outcome, auditors can be cued to more critically evaluate the strength of evidence provided by a weak, aggregate-level analytical procedure based on their existing knowledge of the factors contributing to a procedure's quality.

[1] The term "precision" refers to the quality of the expectation developed by an analytical procedure (SAS No. 56, AICPA 1988). Precision is comprised of four factors: reliability of underlying data, level of aggregation of underlying data, type of analytical procedure model used (e.g., ratio, trend, reasonableness), and inherent predictability of the account or target of the expectation.

Although prior studies examine how explicit instructions (Kinney and McDaniel 1995) and consideration of source competence (Anderson and Koonce 1995) can improve the expectations derived from analytical procedures, none of these studies examines how differences in outcome (i.e., whether or not the difference is significant) between the auditor's expectation and the client's unaudited number can affect the perceived reliability of the analytical procedure itself. Nor do prior studies examine whether an *ex ante* cue might be effective in reducing auditors' tendency to over-estimate the strength and quality of an aggregate-level analytical procedure.

In the first experiment, we manipulate between participants (practicing audit seniors) whether the expected interest revenue derived from an imprecise substantive analytical procedure differs significantly from the client's unaudited number (i.e., whether the low-quality procedure indicates a significant difference or no significant difference). By subsequently providing additional information and having all participants calculate a more precise analytical procedure that reveals the true underlying state of nature (material misstatement), we also manipulate within-subjects the aggregation level, and thereby the precision, of the analytical procedure.[2]

Though the *quality of the evidence* provided by the substantive analytical procedure is identical in both initial outcome conditions (significant difference versus no significant difference), results indicate that, all else equal, when the expected interest revenue derived from the weak initial analytical procedure does *not* differ significantly from the client's unaudited number (i.e., a "favorable" outcome), auditors judge the quality of the evidence obtained from the procedure to be significantly higher than when the expected interest

revenue differs significantly from the client's unaudited number. Moreover, after completing the second, relatively disaggregated procedure, auditors in the no-significant-difference (i.e., "favorable") condition significantly revise downward their initial assessments of the evidence quality they had previously attributed to the highly aggregated analytical procedure, while auditors in the significant-difference (i.e., "unfavorable") condition do not alter their assessments of the initial procedure.

Notably, after receiving additional information and calculating and assessing the second, more precise, analytical procedure, auditors from both outcome conditions reassessed the strength of the initial analytical procedure at a similarly low level. The level of these reassessments is not significantly different from the initial assessment provided by auditors in the significant-difference condition, suggesting that this group perceived the weakness of the initial aggregate-level analytical procedure *before* they observed the relatively disaggregated analytical procedure. Because a "favorable outcome" bias is by definition only operational in the initial assessment of the weak analytical procedure in the no-significant-difference condition, the other condition and the re-assessment measures function as control groups.

Together, these results provide evidence of factors that could lead to over-reliance on weak substantive analytical procedures. Namely, when weak, high-level substantive analytical procedures yield no significant difference, auditors may be insensitive to the imprecision of the expectation and may overestimate the evidential strength provided by the procedure. These results may partly explain the POB finding that auditors often inappropriately rely on weak analytical procedures as substantive evidence supporting a clean audit opinion. When the initial, low-quality procedure indicated no significant difference, the full impact of the

[2] The term "aggregation" simply refers to the level of detail in the underlying data used to form an expectation (e.g., yearly is more aggregated than monthly, entity-level data is more aggregated than business segment data,

imprecision of the weak aggregate procedure was not reflected in auditors' assessments, but was reflected in their re-assessments after they had experienced the results provided by the more precise disaggregated procedure.

In the second experiment, we examine an intervention meant to sensitize auditors to the relatively low quality evidence provided by the imprecise substantive analytical procedures. Although our first experiment suggests that auditors may understand how the precision of a substantive analytical procedure affects the evidential strength derived from that procedure, our results indicate that auditors may overlook a lack of precision if the procedure provides an expectation that is not significantly different from the unaudited values. To examine whether auditors can be cued to consider the precision of a weak analytical procedure *ex ante*, we manipulate between subjects whether auditors receive explicit instruction to consider the weaknesses of an imprecise substantive analytical procedure prior to considering the results yielded by the procedure. We also manipulate the experience level of the participants by including both experienced auditors and students in a graduate financial statement auditing course as participants.

Results from our second experiment indicate that auditors asked to explicitly describe the potential weaknesses in the aggregate procedure prior to calculating its results attribute significantly less evidential strength to that procedure than auditors not explicitly asked to consider its weaknesses. These results suggest that experienced auditors have relevant knowledge that can be cued *ex ante*, and that an intervention in which auditors explicitly consider the weaknesses of an analytical procedure may reduce (but not necessarily eliminate) unwarranted reliance on weak procedures. In contrast, we find that auditing students asked to describe the weaknesses in the aggregate procedure prior to calculating its results *do not*

---

business segment data is more aggregated than individual product data, etc.).

attribute less strength to that procedure than students *not* asked to describe the procedure's weaknesses, suggesting that the intervention depends on a working knowledge of the determinants of precision that auditors possess, but that students do not. Together, these results suggest that auditors have the relevant knowledge to evaluate precision, but that it may be necessary to explicitly cue that knowledge if it is to be applied to mitigate the effects of a bias relating to a favorable outcome.

II. DEVELOPMENT OF HYPOTHESES

Analytical procedures play an important role in a risk-based audit approach. Prior research suggests such procedures can be effective and efficient (Kinney 1987; Wright and Ashton 1989). Analytical procedures also direct attention to high-risk areas, help identify audit issues that detailed work often does not reveal, assist in the evaluation of audit conclusions, and provide substantive evidence (Hirst and Koonce 1996). SAS No. 56 (AICPA 1988) requires the use of analytical procedures at the planning and completion stages of the audit. Auditors perform analytical procedures (e.g., comparisons of unaudited financial data with expected results) at planning to better understand the business and to assist in planning the nature, timing, and extent of testing. At completion, analytical procedures are conducted to ensure that the overall financial statement presentation is consistent with the audit results and the engagement team's knowledge of the business. Because analytical procedures conducted at planning and completion are not primarily focused on substantive evidence, these analytical procedures are often conducted at a high level (e.g., financial statement or business unit level).

Substantive analytical procedures, on the other hand, are conducted in the execution phase of an audit to either indicate the presence of a potential misstatement or to provide

reliable evidence that recorded amounts are free from material misstatement. When analytical procedures are conducted to provide substantive evidence, they must be designed to provide an appropriate level of assurance, taking into account such factors as inherent and control risk, relevant evidence from other sources, etc. Thus, in performing substantive analytical procedures, the auditor must give detailed attention to underlying relationships and develop a precise and independent expectation of the account balance or other financial statement component. The potential effectiveness of an analytical procedure and the degree of reliance that can be placed on the procedure is affected by the quality of the expectation that is developed. The closeness of an expectation to the "correct" amount is called the degree of *precision*. SAS No. 56 (AICPA 1988) suggests that one way to increase the precision of an expectation is to disaggregate the inputs:

> *"Generally, the risk that material misstatement could be obscured by offsetting factors increases as a client's operations become more complex and more diversified. Disaggregation helps reduce this risk .Expectations developed at a detailed level generally have a greater chance of detecting misstatement of a given amount than do broad comparisons. Monthly amounts will generally be more effective than annual amounts and comparisons by location or line of business usually will be more effective than company-wide comparisons.*

In the experimental setting, for example, disaggregating an annual combined loan balance into quarterly loan balances by loan category (e.g., commercial and real estate loans) increases the precision of the resulting expectation for interest revenue.

As mentioned above, the POB's Panel on Audit Effectiveness (2000) found that when substantive analytical procedures are used as the primary test (i.e., the procedure provides the primary assurance for a particular financial statement assertion), in about 20 percent of the cases the analytical procedures used are insufficient to provide the necessary level of assurance.[3] Over-reliance on a weak analytical procedure can affect audit effectiveness and

[3] The Panel (2000) found that analytical procedures are used as the primary or sole substantive test in about 25% of the audit areas. There is evidence that analytical procedures are expected to play an ever-increasing role in

efficiency. That is, a high-level analytical procedure based on imprecise expectations may not yield a significant difference when, in fact, there is a material misstatement in the underlying detail (audit effectiveness).[4] Or, high-level analytical procedures may yield a significant difference when, in fact, there is no material difference, leading to unnecessary investigation (audit efficiency).[5]

Precision of Substantive Analytical Procedures and Outcome-Contingent Judgment

Auditors may over-rely on weak substantive analytical procedures either because they do not understand or they are not sensitive to the low evidence quality provided by weak, aggregate analytical procedures under some circumstances.[6] Kinney and McDaniel (1995, 1996) and Blocher and Patterson (1996) argue that precision is the primary determinant of the assurance provided by a substantive analytical procedure. The greater the required level of assurance, the more precise the expectation must be. If auditors do not understand the relationship between precision and evidence quality, or if they do not think about precision in the normal course of evaluating the strength of evidence provided by analytical procedures, they may be prone to inappropriately over-rely on weak procedures that yield no significant difference. Further, even if auditors understand the importance of precision to the quality of analytical procedures, their sensitivity to precision may depend on other factors. For example, McDaniel and Simmons (2003) find that auditors distinguish between expectations formed

---

future audit approaches (e.g., Bell et al. 1997 and SAS No. 99). Furthermore, analytical procedures are also commonly used as a complimentary substantive test.

[4] Even when a material misstatement doesn't exist and no difference is identified by a weak analytical procedure, over-reliance can occur because the weak analytical procedure would not yield sufficient competent evidence to support the audit opinion.

[5] The use of high-level substantive analytical procedures is common. Our discussions with members of the Auditing Standards Board and practicing auditors, including participants of this study, indicate that the most common form of substantive analytical procedure is a simple 2-period trend analysis, where the unaudited , aggregate value is compared to the prior value.

using aggregated versus disaggregated data when testing a hard account, but not when evaluating an estimate.

Given their experience with low frequencies of material misstatements in client records (Kreutzfeldt and Wallace 1986), auditors usually expect audit procedures to yield no material misstatements. Further, a "favorable" outcome (i.e., one that yields no significant difference), can be seen as consistent with staff auditors' preferences, because in some circumstances such an outcome indicates that no additional work is needed for the audit area in question. When a highly aggregated or otherwise weak analytical procedure yields no significant difference, we expect this "favorable" outcome to dampen auditors' sensitivity to the imprecision of the analytical procedure. In other words, when evaluating the results of weak, aggregated procedures that indicate no material difference, auditors may over-estimate the strength of the evidence provided by those procedures.

The subconscious tendency to make judgments or evaluate the reliability of a procedure based on the degree to which its outcome is commensurate with the goals or desires of the decision maker has been demonstrated in the psychology and accounting literatures, and is referred to as "wishful thinking," "outcome bias," or "motivated reasoning" (e.g., see Elster 1999; Baron and Hershey 1988; Hastie 1984; Pyszczynski and Greenberg 1987; Kunda 1990; Wilks 2002). This research indicates that decision makers often assess evidence as being of higher quality when the evidence obtained is commensurate with their goal-directed efforts. Auditors may similarly attribute greater evidential strength to an analytical procedure that indicates no significant difference than to a procedure that suggests a potential

[6] Prior studies have investigated other potential causes for over-reliance, such as the presence of a plausible but uncorroborated explanation and the lack of an explicit incentive to misstate the financial statements (e.g., Glover et al. 2000).

misstatement, given a preference on the part of auditors for a no-difference outcome.[7] If auditors are positively biased in their assessment of the quality of an analytical procedure indicating no significant difference, they will be more likely to inappropriately rely on that evidence than if they assessed the quality of the evidence in an objective manner.

Research on inherited hypotheses suggests additional reasons auditors may attribute greater evidential strength to a weak procedure when it does not indicate a significant difference. Given an inherited hypothesis of no significant difference (e.g., no material difference identified in the prior year's audit), this literature suggests that auditors may underestimate the likelihood of alternative explanations (Fischhoff et al. 1978; Mehle et al. 1981; Mehle 1982) and become overconfident in the veracity of the inherited hypothesis (Bedard and Biggs 1991; Heiman 1990; Koonce 1992). Such a non-error expectation may also interfere with the consideration of alternative error and non-error hypotheses (Anderson et al. 1992; Frederick 1991; Moser 1989; Libby and Frederick 1990). Notably, prior research also suggests that overconfidence in inherited hypotheses can elevate the perceived informativeness of confirmatory evidence (Swann and Giuliano 1987).

To test whether auditors understand and are sensitive to precision in an analytical procedure and whether their judgments of evidence strength are outcome-contingent, we ask auditor participants to evaluate two different analytical procedures to evaluate the fairness of interest revenue for the same client. All auditors first conduct an aggregate-level procedure

[7] To provide support for the assumption that auditors prefer a no-difference outcome, we asked 68 auditors (of the same experience level and attending different sessions of the same training program as other participants included in this study) the following question, "When performing analytical procedures, would an auditor rather observe no significant difference or a significant difference? In other words, after computing the difference between the current period value and an expectation (e.g., prior period value or a value computed by a reasonableness test), would an auditor rather find a difference that is within or outside of the scope threshold?" Fifty-six auditors answered they would rather observe a no-difference outcome, nine answered they would rather observe a difference outcome, and three did not provide usable responses. On evaluating the auditors' written

(i.e., based on annual data) and then a disaggregated procedure (i.e., based on quarterly data) on the client's interest revenue account. The first, imprecise procedure indicates either material difference or no material difference, producing a between-subjects "outcome" manipulation in the initial, aggregate-level analytical procedure. The second, relatively precise procedure reveals the same underlying state of nature—material misstatement—for all participants, producing a within-subjects "precision" manipulation.

After completing the second (disaggregate-level) procedure, auditors are asked to reassess the evidential strength of the first (aggregate-level) procedure. If auditors are sensitive to the precision of the aggregate-level procedure when they first evaluate its evidential strength, they would not be expected to lower their estimate of evidential strength after being exposed to the disaggregate procedure. But, if auditors are not sensitive to precision when they first evaluate the strength of the aggregate procedure, they would be expected to lower their estimates of the evidential strength of the aggregate procedure after considering the disaggregated data (which should strongly cue auditors' attention to precision).

Consistent with the idea that auditors understand the implications of imprecision but that outcome-contingent processing may diminish their sensitivity to its impact, we predict that when the expected interest revenue derived from the weak, aggregated analytical procedure does not suggest a material misstatement (i.e., "favorable" outcome), auditors in our study will judge the evidential strength of the procedure to be higher than when the results of the same procedure suggest a potential misstatement. This expectation is summarized in the following hypothesis:

---

explanations, we found the predominant explanation to be that a no-difference outcome indicates less required work.

H1: The assessed strength of audit evidence obtained from the aggregate analytical procedure will be greater for auditors who observe no significant difference from that procedure than for auditors who observe a significant difference.

The reasoning above also suggests that, after calculating and assessing the second procedure, auditors in the "no significant difference" condition will reduce their assessment of the first procedure's evidential strength more than will auditors in the "significant difference" condition. This expectation is summarized in the following hypothesis:

H2: After evaluating the disaggregated procedure, auditors who initially observe no significant difference will lower their assessments of the aggregate procedure's evidential strength significantly *more* than will auditors who initially observe a significant difference.

The main and interaction effects predicted by H1 and H2 are depicted in Figure 1, Panel A.

Ex ante Cueing and Relevant Task Knowledge

If auditors understand the importance of precision in evaluating analytical procedures, but are insensitive to imprecision when weak procedures provide results indicating no material difference, the question remains as to whether they can be cued to increase their sensitivity to imprecision *ex ante*. Research in psychology suggests that explicit instructions may help decision makers focus on alternative explanations to a given finding. Petty and Cacioppo (1986), for example, propose an elaboration continuum in decision-makers' evaluations of explanations. This continuum ranges from no thought about relevant information to intense, objective scrutiny of supporting arguments. *Ex ante* warnings, such as an explicit cue to consider potential weakness of an analytical procedure, may lead an auditor to more carefully scrutinize an aggregate-level procedure and to develop counter-arguments to explain positive outcomes (i.e., outcomes indicating no significant difference). This form of

critical information processing can facilitate appropriate rejection of an inherited hypothesis (Gilbert 1991).

Successful *ex ante* cuing depends critically on whether the cued decision maker possesses relevant task knowledge and understanding available for cuing. A cue should impact decision makers with relevant task knowledge and understanding, but should have little or no impact on decision makers that lack such knowledge and understanding. In experiment 2, we examine whether auditors' task knowledge relating to precision can be cued to sensitize them to the implications of imprecise analytical procedures *ex ante*. Experiment 2 also includes participants with a general understanding of auditing concepts but who lack substantial task knowledge or experience. These latter subjects are graduate accounting students, all of whom had previously completed a one-semester undergraduate course in auditing.[8]

Including participants that differ in relevant task knowledge strengthens our ability to investigate whether the cue is activating knowledge which affects judgment.[9] If the effectiveness of the cue depends on the availability of knowledge about the link between the precision of an analytical procedure and evidential strength, auditors should be more affected by an *ex ante* cue than auditing students, who have less relevant knowledge.

Consistent with prior research suggesting that cuing can be effective in increasing objective scrutiny of alternative arguments if requisite task knowledge is available, we expect that an *ex ante* cue will sensitize auditors to the limitations of weak, aggregate analytical

[8] These students have a good overall understanding of the fundamentals of auditing, but they do not have experience in applying analytical procedures in practice, and they lack training on the factors that determine the precision of analytical procedures.

[9] This design choice also allows us to rule out the possibility of a demand affect—i.e., that all cued subjects will reduce their assessments of evidence strength. If cuing participants to consider the weaknesses of a procedure simply makes them more skeptical of a procedure's quality, we would expect both auditors and auditing students who are cued to become more skeptical of the procedure's strength.

procedures. We expect that the same cue will not have an impact on the evidence quality assessments of auditing students, who do not possess the same level of task knowledge and understanding. This reasoning suggests the following interaction hypothesis:

> H3: The perceived strength of audit evidence obtained from a weak, high-level analytical procedure will be lower for auditors who are cued to critically evaluate the quality of the analytical procedure *ex ante* than for auditors who are not cued to critically evaluate the procedure, but the perceived strength of audit evidence will not be lower for auditing students who are cued than for students who are not cued.

III. EXPERIMENT 1

Method

Participants

Sixty-seven supervising seniors with an average of 34.5 months of audit experience (standard deviation 12.8) participated in experiment 1. We chose supervising seniors as participants because they are typically charged with performing analytical procedures (Prawitt 1995; Hirst and Koonce 1996) and have a reasonably well-developed knowledge of potential causes of unexpected fluctuations (Libby and Frederick 1990). Supervising seniors are also directly involved in most aspects of the audit from planning to final review. Participants completed the case materials as part of training sponsored by their firm.[10]

Materials

The experiment asked participants to evaluate the reasonableness of loan interest income for a hypothetical bank. Participants were told that because controls were considered strong, the primary source of substantive evidence regarding the fairness of loan interest

[10] Two of the 67 participants failed to provide an assessment of evidential strength in Part 1 of the experiment, so our analyses present results for only the remaining 65 participants. No significant differences were found in

income would come from substantive analytical procedures, and that additional detail testing would not be performed if analytical procedures provided adequate evidence that interest income is not materially misstated. The case indicated that a misstatement of $525,000 was to be considered material (approximately 4.2% of net income).

Design

The experiment employs both a within-subjects and a between-subjects manipulation. The within-subjects manipulation involves the precision of the analytical procedure. As described previously, all participants were first asked to calculate expected annual interest income for a hypothetical bank using *annual* loan receivable balances and a weighted average *annual* interest rate for all loan categories combined. After doing this and making relevant assessments, all participants were then asked to calculate expected annual interest income using *quarterly* loan receivable balances and weighted average *quarterly* interest rates by loan category (i.e., commercial, real estate, individual and other). This manipulation was meant to test auditors' sensitivity to the precision of the analytical procedure and serve as a cue for participants to objectively reconsider the quality of the first (aggregate) procedure.

The between-subjects manipulation pertained to the outcome of the *first* analytical procedure (i.e., the analytical procedure based on *annual* aggregate loans receivable). Participants were randomly assigned to one of two conditions. For half of the participants, the aggregate-level analytical procedure yielded an expected annual interest income amount that differed materially ($792,300) from that reported by the client. For the other half of the participants, the *same* analytical procedure yielded no material difference ($146,200, "favorable" outcome). To manipulate whether the aggregate procedure indicated the presence

---

participant experience levels between experimental conditions, or in evidence quality assessments between training sessions within experimental conditions.

or absence of a material difference for the annual data while holding constant both the procedure and the underlying quarterly data, we switched some of the quarterly data (e.g., the third quarter data for "Commercial and Agricultural" loans was switched to the fourth quarter) depicted for the disaggregated procedure, thus changing the beginning and end-of-year data to be included in the aggregated procedure while holding the underlying annual data constant in total. Otherwise, all participants saw identical information when calculating the disaggregated analytical procedure. The net aggregate information for the year was the same between conditions. This manipulation allows us to test for the effect of outcome on evidence quality judgments while using an identical analytical procedure between conditions.

Procedures

The experiment was conducted in two Parts. In Part A, participants were provided a copy of the prior year analysis, where an expectation for loan interest income was developed using average *annual* loan volume multiplied by the weighted average *annual* interest rate. Participants were then provided with information to conduct a similar analysis in the current year. Only beginning- and end-of-year information was presented in Part A, as inputs to the *aggregate-level* procedure. Based on the results of the analytical procedure, participants made three assessments. First, participants indicated whether they could accept interest income as reported (yes or no). Next, participants indicated the likelihood that interest income was materially misstated on a scale from 0 to 100, where 100 was labeled "*definitely misstated*." Finally, participants evaluated the "*strength (quality and sufficiency) of evidence provided by the interest income analytical procedure*" on a 7-point scale, with the low point labeled, "*Extremely Weak/Useless Evidence*" and the high point labeled "*Extremely Strong/Removes all Doubt*."

After submitting Part A, participants completed Part B. In Part B, participants used *quarterly* interest rates and *quarterly* loan balances separated by loan category to calculate a new interest income expectation. Because it was based on the same underlying quarterly data, the outcome of this second, disaggregated analytical procedure was the same for all participants—the procedure indicated a potential material misstatement in interest income.[11] After calculating this new interest income expectation, participants answered the same three questions as in Part A, plus the following additional question:

> *"Now reevaluate the analytical procedure used in Part A. Please indicate on the scale below the strength (quality and sufficiency) of evidence provided by the interest income analytical procedure in Part A (average loan volume by average interest rate)."*

This question was answered on the same 7-point scale as before, with endpoints labeled "*Extremely Weak/Useless Evidence*" and "*Extremely Strong/Removes all Doubt*."

Results

The main and interaction effects observed in experiment 1 are depicted in panel B, Figure 1. A visual comparison of panel A and panel B of Figure 2 indicates the observed judgments in experiment 1 are consistent with our expectations. To statistically examine H1 and H2, we conduct a repeated-measures ANOVA on auditors' initial assessments and reassessments of the evidential strength of the aggregate analytical procedure. The means and planned contrasts of this analysis are reported in Table 1. H1 posits that the assessed strength of audit evidence obtained from the aggregated analytical procedure is higher for auditors who observe a "favorable" outcome (i.e., no significant difference) than for auditors who observe a significant difference. Based on the planned contrast reported in Table 1, Panel B,

[11] In this paper, we focus on a setting, or "state of nature," where there appears to be a material misstatement because this is the setting where over-reliance on a weak analytical procedure is of most interest to regulators

we find a significant main effect for outcome (p<0.001), supporting H1. Specifically, auditors whose aggregate analytical procedure indicates no significant difference attribute significantly more evidential strength to that procedure (3.9) than do auditors whose identical aggregate analytical procedure suggests a potential misstatement (2.9). It is important to note that an assessment of the strength of the evidence is not normatively dependent on the outcome of the test. Auditors in both conditions had sufficient information to determine the strength of the evidence. Absent a "favorable" outcome bias, we would expect to observe similar evidential strength ratings across conditions. This result suggests that a "favorable" outcome resulted in a higher assessment of the evidential strength of weak analytical procedures.

H2 predicts that after observing the disaggregated procedure, auditors in the "favorable" outcome condition will lower their previous estimates of the perceived evidential strength of the aggregate procedure significantly *more* than auditors in the significant-difference condition. Auditors' reassessments of the aggregate analytical procedure after seeing the more precise procedure (which reveals the true state of nature—i.e., material misstatement), should be relatively clear and objective. Thus, the difference between the original assessment and the *ex post* reassessment can be seen as a within-subjects correction of the initial assessment.

To test H2, we analyze whether the change in assessments of the aggregate procedure's evidential strength depends on the outcome of the initial aggregate procedure. Using the repeated-measures ANOVA described above, we examine whether the change between the first assessment and the reassessment differs between outcome conditions. As reported in Table 1, Panel B, we find that the degree to which auditors lower their assessments of the aggregate procedure's evidential strength strongly depends on the outcome

---

and users of financial statements.

of the first (aggregate) analytical procedure, supporting H2 (p<.001). Specifically, auditors initially observing a significant-difference outcome from the aggregate procedure do not significantly lower their original assessments of that procedure's evidential strength (2.9 and 2.8, respectively, t_{35}=0.23, p=0.81); however, auditors initially observing a no-significant-difference outcome from the aggregate procedure do significantly lower their assessments (from 3.9 to 2.7, t_{28}=6.12, p<0.001). Notably, the final reassessment of the aggregate analytical procedure's quality does not differ between participants in the two outcome conditions (2.8 vs. 2.7, t_{63}=0.56, p=0.58).

An examination of Figure 3, Panel B, highlights that the evidential quality assessments in three of the four conditions are statistically identical; the only condition producing a result that is different from the other three is the Part A, no-significant-difference (favorable outcome) condition. The other three experimental cells are reasonably viewed as being representative of unbiased judgments, especially in consideration of the reassessment scores. Using these three cells as a basis for comparison, it appears that auditors in the Part A, "favorable outcome" condition overestimated the strength of evidence provided by the weak analytical procedure.

In evaluating our H1 and H2 results, it is important to understand that 1) outcome is the only difference between the conditions, and 2) the strength of evidence provided by the weak, high-level analytical procedure in Part A does not normatively depend on the outcome. Rather, the strength of the evidence depends on the precision of the expectation developed. Auditing standards require sufficient substantive audit evidence for all significant accounts, regardless of the auditor's planned reliance on controls. The weak analytical procedure in this study does not produce strong evidence because the expectation is insufficiently precise.

Considering our H1 and H2 results together, we find that when a weak, aggregate-level procedure produces a significant-difference outcome or when auditors can compare a weak, aggregate-level procedure to a disaggregated procedure, they assess quality of the weak analytical procedure as low. An outcome that conflicts with beliefs or preferences (i.e., significant difference) likely serves as a cue to more objectively evaluate the quality of the analytical procedure that produced the outcome. This argument is bolstered by the reassessment results for auditors in both conditions. However, when the weak procedure's outcome is consistent with auditors' beliefs or preferences (i.e., "favorable"), they over-estimate the quality of the weak analytical procedure.

To further examine the potential impact of the results relating to H1 and H2, we also asked participants to assess the likelihood that interest income is materially misstated, based on the results of the analytical procedures (on a scale of 0 to 100%, where 100% is labeled, "Definitely Misstated"). In Part A, the average assessed likelihood of material misstatement is significantly higher for the significant-difference group than for the no-significant-difference group, as expected (41% versus 11%, respectively; $p<.001$). Further, the no-significant-difference group's average assessment increased from 11% in Part A to 48% in Part B ($p<.001$). Finally, there is a significant negative correlation between the no-significant-difference group's initial evidential quality assessments and their assessments of the likelihood of material misstatement (Pearson correlation of -0.47; $p<.001$), indicating that as these participants' assessments of evidence quality increase, their assessments of the likelihood of material misstatement decrease. These results are consistent with the idea that participants' evidential quality assessments are meaningful to important subsequent audit judgments as they gather evidence to support their opinion on the fairness of the financial

statements, and that over-assessments of evidence quality may indeed ultimately lead to over-reliance on weak analytical procedures.

IV. EXPERIMENT 2: EX ANTE CUING AND EXPERIENCE

Experiment 1 provides evidence that auditors' assessments of the strength of evidence provided by relatively weak, high-level analytical procedures are outcome-contingent. While the participants in the no-significant-difference condition were able to make *ex post* corrections for their apparent over-reliance in Part A of experiment 1 after conducting an improved analytical procedure based on disaggregated data (Part B), the question remains whether the demonstrated over-estimation of evidence strength can be reduced *ex ante* by cuing experienced auditors' relevant knowledge. We conduct a second experiment to address H3 and test whether an *ex ante* intervention can reduce auditors' tendency to over-estimate the strength of weak analytical procedures that yield no significant difference. In H3 we predict that an *ex ante* cue will sensitize auditors to the limitations of weak, aggregate analytical procedures. We also predict that the same cue will not have an impact on the evidence quality assessments of auditing students, who do not possess the same level of task knowledge and understanding.

Method

Eighty-four supervising seniors with an average of 38.5 months of audit experience (standard deviation 13.6) and ninety inexperienced graduate-level students in an advanced auditing course at a large private university participated in experiment 2.[12] None of the participants in experiment 2 participated in experiment 1. The case materials and related questions are essentially identical to the no-significant-difference condition of Part A,

experiment 1.[13] As in experiment 1, the underlying state of nature is that a material misstatement exists.[14] Holding constant all other information across participants, we manipulate only one variable between subjects. Half of the participants read the following additional paragraph in the instructions prior to calculating the result of the aggregate analytical procedure:

> *Before computing an expectation of interest income for the current year, please consider the strength (quality and sufficiency) of the evidence provided by the interest income analytical procedure used by the audit team last year. Please list in the space provided below one or more weaknesses of this analytical procedure.*

In summary, the experimental design for experiment 2 focuses on a between-subjects manipulation involving an *ex ante* cue to consider the strength of the analytical procedure (i.e., explicitly cued to critically evaluate the prior year aggregated procedure versus not cued to critically evaluate), crossed with an experience manipulation (auditors versus graduate accounting students).

Results

In accordance with our prediction in H3, we compute an ANOVA to examine whether an explicit cue to critically evaluate the aggregate procedure prior to calculating the results affects auditors', but not students', assessments of evidential strength for the aggregate procedure. Based on the means and planned contrasts reported in Table 2, the experience by cuing interaction is significant (p=.011), supporting H3. The contrast for auditor participants between cuing conditions indicates a significant difference in perceived evidential strength as

[12] No significant differences were found in auditor experience levels between experimental conditions, or in evidence quality assessments within experimental condition between training sessions.
[13] In experiment 2, we described the prior year analytical procedure rather than provide it in tabular form.

a result of the *ex ante* intervention (p<0.001). Specifically, auditors who were explicitly asked to consider potential weaknesses of the aggregate analytical procedure attributed less evidential strength to that procedure (3.3) than did auditors who were not asked to consider potential weaknesses (4.1). This suggests that a relatively simple instruction to focus auditors' attention on potential weaknesses may be useful in triggering consideration of the imprecision of weak high-level, aggregate analytical procedures.[15]

However, the contrast for student participants between cuing conditions indicates no significant difference in perceived evidential strength as a result of the cuing manipulation (p=0.64, two-tailed). Specifically, students who were cued to consider potential weaknesses of the aggregate analytical procedure attributed the same level of evidential strength to that procedure (2.8) as students who were not cued (2.9). Thus, a simple instruction to focus students' attention on potential weaknesses did not result in a lower assessment of evidence strength (relative to the un-cued students), which suggests the cue alone is insufficient to reduce evidence strength assessments for students who do not possess the relevant knowledge for the cue to be effective.[16]

---

[14] We focus on the no-significant-difference (i.e., favorable) condition to examine the effect of cuing because this is the condition in which over-estimation of evidence strength occurred in experiment 1, and therefore is the condition where cuing might have a positive effect on auditor judgments.

[15] This result also strengthens the argument that auditors' assessments of evidence quality in the no-significant-difference outcome condition are inappropriately high in both experiments in the absence of cuing. Note, however, that cuing in experiment 2 did not reduce auditors' quality assessments to the same degree as did conducting a disaggregated procedure that contradicted the results of the first procedure, as in experiment 1.

[16] Students' assessments of evidential strength are lower on average than the assessments of auditors, but the difference in average assessment *level* between students and professionals is not of central importance in this study—note that the effect of the cue manipulation between subjects is a *difference* measure. While not plausible, it is possible that, unlike experienced auditors, students saw through the weakness of the procedure without being cued, which could provide an alternative explanation for the inefficacy of the cue for student subjects and their lower assessment that the un-cued auditors. While not tabulated, we did have the students complete Part B of the experiment. Therefore, to address the concern that students did not overestimate the evidential strength in Part A, we compared students' Part A and Part B reassessment responses. Students' average Part A assessment is 2.87 and their average reassessment of Part A, after seeing the relatively disaggregated procedure, is 1.82. This difference is significant at p < .001 (t_{89}=7.35). The cue does not mediate the assessment/reassessment relationship (F_{88}=0.10, p=0.757). Together, these results suggest that the students overestimated the strength of evidence in Part A. The difference in mean assessment *levels* between students and

V. CONCLUSION

Various high-profile frauds provide examples of inappropriate auditor reliance on weak audit evidence. The Panel of the Public Oversight Board charged with investigating audit effectiveness reported evidence in 2000 that auditors frequently over-rely on weak substantive analytical procedures as evidence supporting a clean audit opinion. Using substantive analytical procedures as our context, we examine a partial explanation for why auditors may inappropriately rely on weak audit evidence. In two experiments involving a total of 171 supervising seniors and 90 auditing students, we provide evidence consistent with the idea that insensitivity to the imprecision of weak analytical procedures that yield a "favorable" (i.e., no significant difference) outcome is a potential cause for auditor over-reliance on weak audit evidence. In experiment 1, we find that auditors attributed significantly more evidential strength to a weak analytical procedure that suggested no significant difference than to an identical procedure based on the same underlying data that suggested a potential misstatement. Further, when the auditors in the no-significant-difference condition reconsidered the aggregated analytical procedure after conducting stronger, more precise analytical procedures based on disaggregated data, they significantly reduced their strength-of-evidence ratings for the aggregate analytical procedure, while auditors in the significant-difference condition did not. Notably, both outcome groups reassessed the quality of the initial aggregate-level analytical procedure at the same low level after receiving further information and calculating the disaggregated analytical procedure; this reassessment is at the

---

auditors is not surprising as students do not have experience in making such assessments, and in our experience students tend to be more negative in general than practicing auditors (perhaps because college case studies tend to focus on audit failures).

same level at which auditors in the significant-difference condition initially assessed the aggregated procedure's quality.

In experiment 2, we examined whether the insensitivity to the potential weaknesses in highly aggregated analytical procedures can be reduced *ex ante* by explicitly cueing participants to critically evaluate the aggregate analytical procedure prior to executing it, and whether the effect of cueing would differ between professional and student participants. In experiment 2, half the participants were asked to critically evaluate the aggregated analytical procedure performed in the prior year before conducting the current year analytical procedure. We found that the explicit consideration of potential weaknesses associated with the highly aggregated analytical procedure conducted in the prior year did result in increased sensitivity to the imprecision in the current-year procedure on the part of auditors, but that the cue had no significant effect on students' assessments. We argue that the effectiveness of the *ex ante* cue in experiment 2 depends on auditors' understanding of the link between the precision of an analytical procedure and evidence strength. Because accounting students lack such an understanding (or have a much less developed understanding) the guidance does not significantly affect their judgments, strengthening the conclusion that relevant task knowledge is the construct being triggered by the cue, and helping to rule out the possibility that the cue simply created a demand effect.

This study is subject to the usual limitations of behavioral auditing research. The generalizability of our results may be limited by the fact that participants came from one big-4 accounting firm. However, the POB report did not indicate differences in the degree of over-reliance on weak analytical procedures between firms. Participants in our study worked alone and thus we do not measure the potential mitigating effects of teamwork or the review process

that pervade the audit environment. However, the POB report was based on an examination of actual audit files; thus, it appears that the common working practices of audit firms (e.g., review processes) did not remove the over-reliance on weak analytical procedures observed by the POB.

REFERENCES

American Institute of Certified Public Accountants. 1988. Statements on Auditing Standards No. 56. *Analytical Procedures*. New York, NY: AICPA.

American Institute of Certified Public Accountants. 2002. Statements on Auditing Standards No. 99. *Consideration of Fraud in a Financial Statement Audit*. New York, NY: AICPA.

Ameen, E., and J. Strawser. 1994. Investigating the use of analytical procedures: An update and extension. *Auditing: A Journal of Practice and Theory* (Fall): 69-76.

Anderson, J. C., S. E. Kaplan and P. M. J. Reckers. 1992. The effects of output interference on analytical procedures judgments. *Auditing: A Journal of Practice and Theory* (Fall): 1-13.

Anderson, U., and L. Koonce. 1995. Explanation as a method for evaluating client-suggested causes in analytical procedures. *Auditing: A Journal of Practice and Theory* (Fall): 124-132.

Arens, A., and J. Loebbecke. 2000. *Auditing. An Integrated Approach*. Englewood Cliffs, NJ: Prentice Hall.

Bamber, M. E. 1983. Expert judgment in the audit team: A source reliability approach. *Journal of Accounting Research* 21 (Fall): 396-413.

Baron, J., and J. Hershey. 1988. Outcome bias in decision evaluation. *Journal of Personality and Social Psychology* 54 (4): 569-579.

Bedard, J., and S. Biggs. 1991. The effects of domain-specific experience on evaluation of management representations in analytical procedures. *Auditing: A Journal of Practice and Theory* (Supplement): 77-95.

Bell, T., F. Marrs, I. Solomon, and H. Thomas. 1997. *Auditing Organizations Through a Strategic-Systems Lens: The KPMG Business Measurement Process.* New Jersey: KPMG Peat Marwick LLP.

Blocher, E. and G. Patterson, Jr. 1996. The use of analytical procedures. *Journal of Accountancy.* Volume 181 (February): 53-55.

Caster, P., and K. Pincus. 1996. An empirical test of Bentham's theory of the persuasiveness of evidence. *Auditing: A Journal of Theory and Practice* 15 (Supplement): 1-22.

Elster, J., 1999. *Nuts and Bolts for the Social Sciences*. Cambridge University Press.

Fischhoff, B., P. Slovic, and S. Lichtenstein. 1978. Fault trees: sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance* (May): 330-344.

Frederick, D. M. 1991. Auditors' representation and retrieval of internal control knowledge. *The Accounting Review* (April): 464-485.

Gilbert, D. 1991. How mental systems believe. *American Psychologist* (February): 107-19.

Glover, S., J. Jiambalvo and J. Kennedy. 2000. Analytical procedures and audit planning decisions. *Auditing: A Journal of Practice and Theory,* Vol. 19, No. 2, (fall).

Hastie, R. 1984. Causes and effects of causal attribution. *Journal of Personality and Social Psychology* 46: 44-56.

Heiman V. B. 1990. Auditor's assessments of the likelihood of error explanations in analytical review. *The Accounting Review* (October): 875-891.

Hirst, E. 1994. Auditors' sensitivity to earnings management. *Contemporary Accounting Research* (Fall): 405-422.

Hirst, E., and L. Koonce. 1996. Audit analytical procedures: a field investigation. *Contemporary Accounting Research* 13 (2): 457-486.

Kinney, W. Jr. 1987. Attention-directing analytical reviews using accounting ratios: a case study. *Auditing: A Journal of Practice and Theory* (Spring): 59-73.

Kinney, W. Jr and L. McDaniel. 1995 Expectation-formation guidance in the auditor's review of interim financial information, *Journal of Accounting Research,* Vol. 33 (spring):33-

Kinney, W. Jr and L. McDaniel. 1996. How to improve effectiveness of substantive analytical procedures. *The CPA Journal,* Vol. 66 (April): 52-

Koonce. L. 1992. Explanation and counterexplanation during audit analytical review. *The Accounting Review* (January): 59-77.

Koonce, L., U. Anderson, and G. Marchant. 1995. Justification of decisions in auditing. *Journal of Accounting Research* (Autumn): 369-385.

Kreutzfeldt, R., and W. Wallace. 1986. Error characteristics in audit populations: their profile and relationship to environmental factors. *Auditing: A Journal of Practice and Theory* (Fall): 20-43.

Kunda, Z. 1990. The case for motivated reasoning. *Psychological Bulletin* 108 (3): 480-498.

Libby, R., and D. Frederick. 1990. Experience and the ability to explain audit findings. *Journal of Accounting Research* (Autumn): 348-367.

McDaniel, L., and L.E. Simmons. 2003. Testing estimates with analytical procedures: Auditors' Evaluations of Expectation Precision. Working paper, University of Kentucky.

Mehle, T. 1982. Hypothesis generation in an automobile malfunction inference task. *Acta Psychologica* (November): 87-106.

_____, C. Gettys, C. Manning, S. Baca, and S. Fisher. 1981. The availability explanation of excessive plausibility assessments. *Acta Psychologica* (November): 127-140.

Moser, D. 1989. The effects of output interference, availability, and accounting information on investors' predictive judgments. *The Accounting Review* (July): 433-448.

Petty, R., and J. Cacioppo. 1986. The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 19). New York: Academic Press: 123-205.

Prawitt, D. 1995. Staffing assignments for judgment-oriented audit tasks: The effects of structured audit technology and environment. *The Accounting Review* 70 (July): 443-465.

Public Oversight Board. 2000. *Report and Recommendations.* The Panel on Audit Effectiveness. Stamford, CT.

Pyszczynski, T., and J. Greenberg. 1987. Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis-testing model. In *Advances in Experimental Social Psychology*, vol. 20, ed. L. Berkowitz, 297-340. New York: Academic Press.

Swann, W., and I. Giuliano. 1987. Confirmatory search strategies in social interaction: how, when, why, and with what consequences? *Journal of Social and Clinical Psychology* (Vol. 5): 511-524.

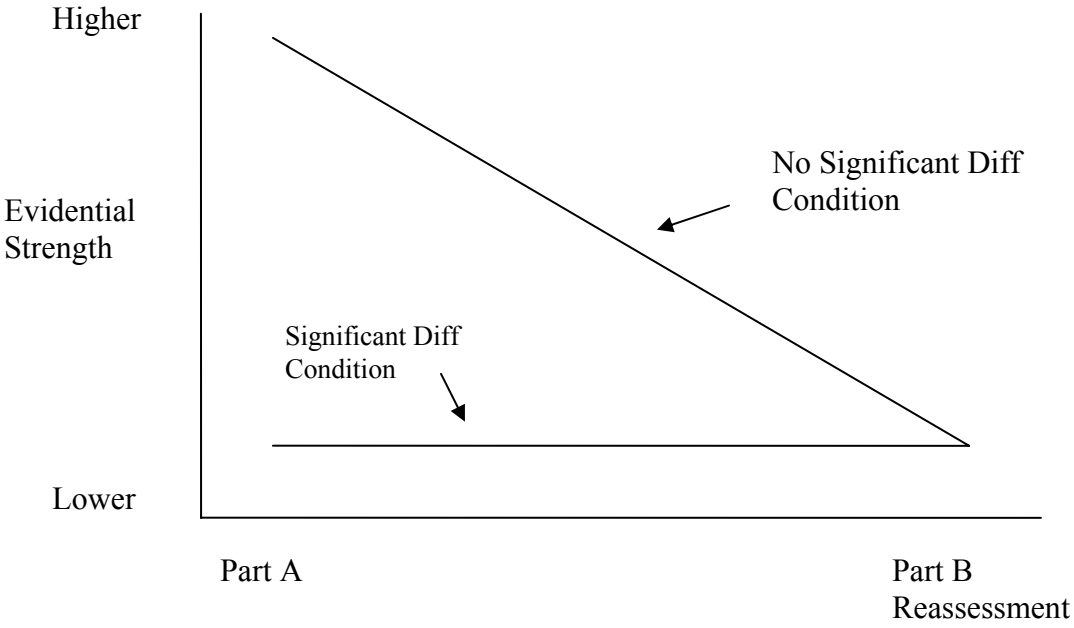
Wilks, T. J. 2002. Predecisional distortion of evidence as a consequence of real-time audit review. *The Accounting Review* (77) 1: 51-71.

Wright, A., and R. H. Ashton. 1989. Identifying audit adjustments with attention-directing procedures. *The Accounting Review* (October): 710-728.

FIGURE 1

Predicted and Observed Effects of Precision and Outcome on Assessment Evidential Strength

Panel A: Illustration of the main and interaction effects predicted by H1 and H2



Panel B: Main and interaction effects observed in Experiment 1

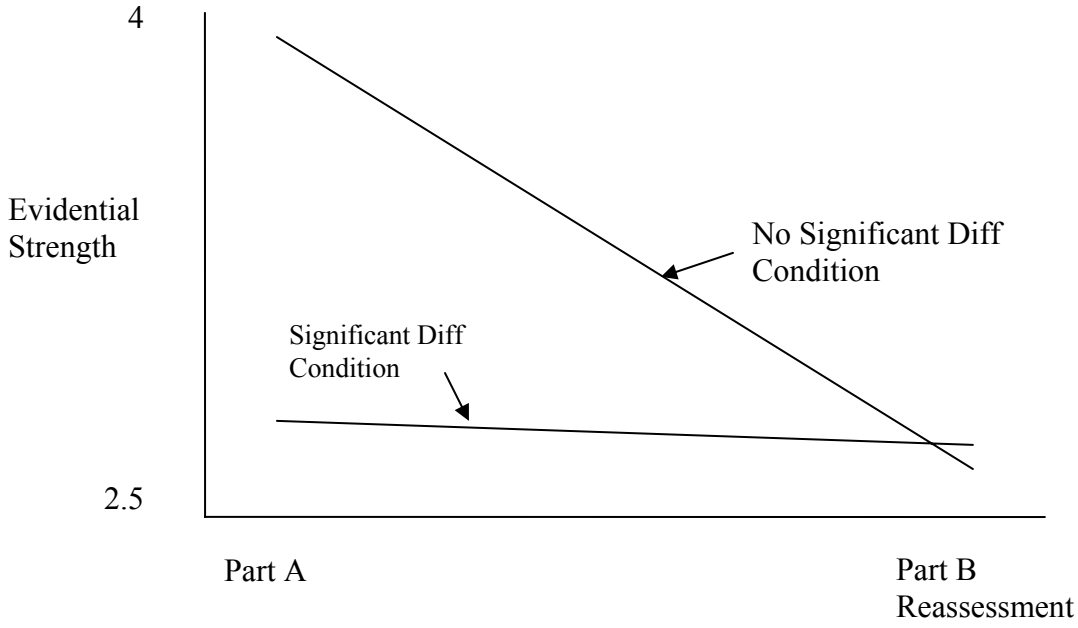


TABLE 1

How Do Precision and Outcome Bias Affect Auditors' Judgments of the Evidential Strength Provided by Analytical Procedures? – Experiment 1

Panel A – Mean Estimates of Evidential Strength[a] (Standard Deviation) [n]

Outcome of *Aggregate* Analytical Procedure	Estimates of Evidential Strength for		
	Part A-Aggregate Procedure (Initial Assessment)	Part B-Disaggregated Procedure – Suggests Material Misstatement	Aggregate Procedure (Reassessment)
No Significant Difference	3.9 (0.81) [29]	3.7 (0.99) [29]	2.7 (1.16) [29]
Significant Difference	2.9 (1.02) [36]	3.8 (1.03) [36]	2.8 (1.21) [36]
Combined	3.4	3.8	2.8

Panel B – Planned Contrasts

		Test Statistic	***p*-Value**[b]
H1:	Effect of analytical procedure outcome on initial assessments of aggregate procedure (between-subjects)	$F_{1,63} = 20.09$	0.0001
H2:	Comparison of change in aggregate-procedure assessments when disaggregated procedure confirms or disconfirms aggregate procedure (within and between-subjects)	$F_{1,63} = 16.02$	0.0001

[a] Assessments of evidential strength were made on a 7-point scale, with the low point labeled "Extremely Weak/Useless Evidence" and the high point labeled "Extremely Strong/Removes all Doubt."

[b] Reported *p*-values are one-tailed.

TABLE 2

How Does Explicit Consideration of Aggregate Procedure Weaknesses Affect Judgments of the Evidential Strength Provided by Analytical Procedures? – Experiment 2

Panel A – Mean Estimates of Evidential Strength[a] (Standard Deviation) [n]

Explicit Guidance to Consider Weaknesses	Participant		
	Professional Participants	Student Participants	Combined
No	4.1 (0.64) [42]	2.9 (1.27) [45]	3.5 (1.18) [87]
Yes	3.3 (1.05) [42]	2.8 (1.13) [45]	3.0 (1.11) [87]
Combined	3.7	2.9	3.3

Panel B – Planned Contrasts

		Test Statistic	*p*-Value[b]
H3:	Experience by Cuing Interaction	$F_{1,170} = 5.34$	0.0110
	Simple Effect of Cuing for Auditors	$F_{1,82} = 13.44$	0.0002
	Simple Effect of Cuing for Students	$F_{1,88} = 0.22$	0.6400

[a] Assessments of evidential strength were made on a 7-point scale, with the low point labeled "Extremely Weak/Useless Evidence" and the high point labeled "Extremely Strong/Removes all Doubt."

[b] Reported *p*-values are one-tailed, except for the simple effect for students because our prediction was non-directional.