

Do Auditors Under-Assess the Severity of Internal Control Design Deficiencies?

Christopher J. Wolfe
Mays Business School
Texas A&M University
4353 TAMU
College Station, TX 77845
(979) 845-0964
Email: cwolfe@mays.tamu.edu

Michelle Chandler Diaz
Mays Business School
Texas A&M University
4353 TAMU
College Station, TX 77845
(979) 845-3275
Email: mdiaz@mays.tamu.edu

August 2009

We sincerely thank the participating firm for providing participants and expert insights into the audit process. We thank Barbara Apostolou, Joe Brazel, Bill Felix, James Flagg, Bill Kinney, Tina Loraas, Elaine Mauldin, Mary Lea McAnally, Mark Peecher, Ed O'Donnell, Tom Omer, Jaime Schmidt, Mike Shaub, Marjorie Shelley, Ed Swanson, Eileen Taylor, and workshop participants at Louisiana State University, Texas A&M University, and the 2008 University of Kansas Audit Symposium for their thoughtful comments. Christopher Wolfe gratefully acknowledges PricewaterhouseCoopers for providing financial support while completing this research.

Do Auditors Under-Assess the Severity of Internal Control Design Deficiencies?

Abstract

Control deviation incidence is not documented for internal control design deficiencies, because controls with deficient designs are not tested for operating effectiveness. Support theory predicts that auditors will not mentally elicit undocumented control deviations and, therefore, will under-assess the severity of design deficiencies. As predicted by support theory, we find that most partners and managers in our experiment's sample under-assess control risk and under-adjust substantive tests for a design deficiency. Senior auditors in our experiment behave theoretically similar to partners and managers, but seniors have difficulty determining the appropriate risk and substantive test to adjust. In summary, our results point to a bias that arises when auditors integrate results from the internal control audit into the financial statement audit. Further, our results suggest that standard audit team review procedures are inadequate to ameliorate this bias.

Key Words: internal control deficiency, risk assessment, substantive testing, support theory

Do Auditors Under-Assess the Severity of Internal Control Design Deficiencies?

1. Introduction

We study auditor risk assessment and substantive test adjustment for internal control design deficiencies. Design deficiencies do not warrant detail control testing, because improperly designed controls are always ineffective and therefore produce maximum control risk (AICPA 2007 ¶ 62-69, ¶ 75-76). Due to the absence of detailed testing, which eliminates the explanation and enumeration of actual control deviations, we theorize that auditors under-assess the severity of design deficiencies. Auditing Standard (AS) 5 requires that deficiencies identified in the internal control audit be integrated into the financial statement audit (PCAOB 2007 ¶ 6-8, B6). A failure to effectively integrate internal control design deficiencies into financial statement audits can increase audit risk and, ultimately, the potential for audit failure (PCAOB 2007 ¶ B4-B8).

Auditors have access to three sources of information when evaluating an internal control deficiency: (1) deficiency documentation, (2) cognitive content about deficiencies, and (3) the ease of mentally retrieving cognitive content about a deficiency. We expect that both evidence content and the ease with which content is retrieved from memory have an effect on deficiency assessment. As such, we draw on behavioral literatures in support theory and memory ease-of-retrieval to develop our expectations.

Support theory indicates that non-normative judgments and decisions occur when evidence about an event is packed into a categorical description (Tversky and Koehler 1994). For example, Brody, Coulter, and Daneshfar (2003) report lower auditor assessment for a misstatement in “some transaction cycle” compared to “investing, financing, payroll, or some other transaction cycle.” The description, “investing, financing, payroll, or some other transaction cycle”, is said to be packed into the categorical description, “some transaction cycle.” Support theory predicts that evidence users do not unpack categorical evidence into its

component parts causing them to make under-assessments (Van Boven and Epley 2003). Applying this concept to control deficiencies, the evidence underlying a design deficiency can be viewed as categorical because it only identifies the nature of the design flaw. The design flaw encapsulates a group of potential control deviations, but specific instances of deviations are unidentified and un-enumerated due to the absence of detailed control testing. Following support theory, the packed evidence contained in design deficiency documentation leads auditors to under-assess control risk and under-adjust planned substantive tests.

Support theory's prediction of design deficiency under-assessment is premised on the nature of deficiency documentation. However, we expect that auditor cognition supplements documented deficiency evidence with (1) a mental picture of the deficiency, and (2) the subconscious perception of the ease of mentally picturing the deficiency (Menon and Raghurir 2003). We expect that easily and clearly picturing a design deficiency produces two effects. First, ease of mentally retrieving a picture of a deficiency should increase the perceived severity of the deficiency (Schwarz and Vaughn 2000). Second, auditors who easily and clearly picture a design deficiency are expected to unpack the categorical evidence of a design flaw. This increases perceived deficiency severity by eliminating support theory bias (Van Boven and Epley 2003). Accordingly, we expect that auditors under-assess the severity of design deficiencies only when they *do not* easily and clearly picture a deficiency.

We investigate deficiency evaluation in an experiment using 87 auditor participants from an international public accounting firm (9 partners, 30 managers, and 48 seniors). The auditors were asked to review an audit case, assess risk, and adjust substantive tests in an audit program. To test our expectations, we benchmark auditor assessments of control risk and substantive test adjustments for a design deficiency against those for an operating deficiency. Operating deficiencies represent execution breakdowns in properly designed controls and are subject to

detail control testing to determine the extent of the execution breakdowns (PCAOB 2007 ¶ 44). Consequently, operating deficiency documentation is unpacked and should not produce a support theory bias, because actual control deviations are named and enumerated. *Ceteris paribus*, support theory predicts that design deficiencies are assessed as less severe than operating deficiencies. This is contrary to professional guidance (AICPA 2007 ¶ 62-76). Design deficiency control risk must always be set at maximum because the control never works, however operating deficiency control risk can be set at less than maximum dependent on detail control testing outcomes.¹ Therefore, a finding that design deficiencies are assessed as less severe than operating deficiencies is consistent with support theory and indicates that design deficiencies are under-assessed.

As expected, we find that partners and managers who *cannot* easily picture the deficiency assess control risk lower and increase substantive tests less for the design deficiency than for the operating deficiency. Seniors behave similar to partners and managers in terms of under-assessing the design deficiency vis-à-vis the operating deficiency, but seniors have difficulty identifying the appropriate risks and substantive procedures to adjust. Contrary to professional guidance, we find that, on average, auditors at all experience levels assess design deficiency control risk at less than maximum. In sum, our results indicate that the bulk of auditors in our study under-assess the severity of the internal control design deficiency.

AS 5 requires that auditors consider the control evaluations from the internal control audit when planning the nature, timing, and extent of substantive tests in the financial statement audit (PCAOB 2007 ¶ B6). From a practical standpoint, our results indicate that seniors under-assess risk and under-adjust substantive tests for design deficiencies. Further, partners and managers,

¹ A system's controls can be classified as "sometimes down" or "always down" in the presence of an operating deficiency; whereas, a system's controls must be classified as "always down" in the presence of a design deficiency (AICPA 2007 ¶ 62-76). Holding transaction cycle constant, a design deficiency's control risk and related increase to substantive tests should be greater than or equal to that of an operating deficiency.

who review senior-level work, make similar risk assessments and substantive test adjustments. From a theoretical standpoint, we document evidence of support theory bias that occurs during the integration of the internal control and financial statement audits. Practically, our results suggest that firms should consider addressing this issue in training for auditors at all ranks.

The remainder of the paper is organized as follows. The next section provides background and develops hypotheses. Then, the experimental methods are described and the results are presented. We conclude with a discussion of our study's implications and limitations.

2. Background and hypotheses

Per AS# 5, auditors are required to analyze control systems in terms of their design and operation (PCAOB 2007 ¶ 42-45). Tests of control system design involve inquiry of client personnel, observation of operations, and inspection of documentation (PCAOB 2007 ¶ 43). A control system that operates according to its design but fails to prevent or detect errors and fraud constitutes a design deficiency (PCAOB 2007 ¶ 42). If a control system is deficiently designed, control testing stops and the system's inability to prevent or detect errors and fraud is documented. A design deficiency is a first-order failure, because the control system can never operate in a preventive or detective fashion. Therefore, a system with a design deficiency is "always down." Control operations are not tested because the controls do not mitigate the related financial statement risk (AICPA 2007 ¶ 62-69). A design deficiency should result in a maximum control risk assessment which precludes reliance on the control system (AICPA 2007 ¶ 62-69).

A control system's operating effectiveness is tested only after its design has been tested and deemed acceptable (PCAOB 2007 ¶ 44). A properly designed control system that operates in a manner that fails to prevent or detect errors and fraud is an operating deficiency (i.e. an operating deficiency exists in a system that has passed a design test). A system with an operating deficiency is only "down" when operating procedures are not being followed. Tests of control

system operation involve inquiry, observation, inspection, and control re-performance and are performed to determine when controls are and are not down (PCAOB 2007 ¶ 45). When testing control system operating effectiveness, auditors observe and document procedural breakdowns that represent actual deviations from control system objectives (PCAOB 2007 ¶ 49-54).

Depending on the outcome of the testing, auditors can assess control risk at less than maximum, thereby placing some reliance on control systems with operating deficiencies (AICPA 2007 ¶ 62-74).

AS# 5 (PCAOB 2007 ¶ B5-B6) requires that identified deficiencies from the audit of internal controls be integrated into the planning of substantive tests in the financial audit. Statement of Auditing Standard (SAS) # 47 requires the use of the audit risk model (ARM) as part of the audit planning process. The ARM is stated as follows (AICPA 1983):

$$\text{Acceptable Audit Risk} = \text{Inherent Risk} \times \text{Control Risk} \times \text{Detection Risk}^2 \quad (1)$$

The ARM is a compensatory model. Inherent risk and control risk are assessed by the auditor, while detection risk is set by the auditor to manage acceptable audit risk. As the auditor assesses control risk higher, detection risk must be reduced to keep audit risk at an acceptable level.

Reducing detection risk is done by increasing substantive testing. Auditing standards provide for increasing substantive testing when control risk is high (AICPA 2007 ¶ 105-108). Because control risk should be set at maximum for systems with design deficiencies, planned substantive tests should be significantly increased for such systems. In contrast, systems with operating deficiencies can have control risk set below maximum and do not necessarily demand significant increases to substantive tests. Systems with an operating deficiency, where testing indicates that

² Acceptable audit risk is the probability that an auditor will give an unqualified opinion on materially misstated financial statements and it is set by the auditor. Inherent risk is the probability that an account balance or class of transactions contains a misstatement in the absence of internal controls and it is assessed by the auditor. Control risk is the probability that error and fraud are not prevented or detected by the internal control system and it is assessed by the auditor. Detection risk is the risk that substantive procedures will not detect a misstatement and it is determined by the nature, timing, and extent of substantive tests.

controls work correctly some of the time, should have lower assessed control risk and a smaller increase to substantive tests than systems with a design deficiency.

Deficiency evidence and auditor judgment

When evaluating a deficiency, auditors have access to three sources of information. The first is deficiency documentation from client-provided and auditor-gathered information. The second is content from the auditor's mental cognition regarding the deficiency. The third is the ease or difficulty of creating cognition for the deficiency (Caruso 2008). Psychology research identifies independent judgment effects for information content and the ease with which information is retrieved from memory (see Schwarz and Vaughn (2000) for a review). We expect that both the contents of evidence and the ease with which evidence is mentally retrieved influence auditor evaluation of control deficiencies. Therefore, we appeal to two theories: support theory and theory on the ease of retrieval from memory.

Support theory informs judgment and decision making that is based in packed versus unpacked evidence (Tversky and Koehler 1994). Specifically, this theory indicates that the support for an event, described as a multifaceted category, increases when the event's description is unpacked into the category's component parts (Tversky and Koehler 1994; Rottenstreich and Tversky 1997). For example, Van Boven and Epley (2003) find that mock jurors assess penalties for pollution that results in "all varieties of respiratory diseases" significantly lower than penalties for pollution that results in "asthma, lung cancer, throat cancer, and all other varieties of respiratory diseases" (e.g., packed and unpacked victim compensation means are \$130,336 and \$264,722, respectively).³

³ A number of psychological studies confirm expected effects of evidence packing/unpacking (Fishcoff, Slovic, and Lichtenstein 1978; Fox and Birke 2002; Kruger and Evans 2004; Johnson, Hershey, Meszaros, and Kunreuther 1993; Macchi, Osherson, and Krantz 1999; Redlmeier, Koehler, Liberman, Tversky 1995; Rottenstreich and Tversky 1997; Russo and Kolzow 1994; Savitsky, Van Boven, Epley, and Wight 2005; Tenney, Cleary, and Spellman 2009; Tversky and Koehler 1994; Van Boven and Epley 2003).

We transfer support theory's definition of packed and unpacked evidence to control deficiencies as follows. For design deficiencies, auditors document only the nature of the control system's design flaw. Evidence documenting design deficiencies is *packed*, because transaction-level deviations are not identified, rather they are encapsulated in the definition of the design flaw. In contrast, auditors document the types and frequency of transaction-level deviations associated with operating deficiencies. Evidence documenting operating deficiencies is *unpacked*, because transaction-level deviations are identified and enumerated. In sum, design deficiency documentation indicates that "some group of transaction-level deviations exists" and operating deficiency documentation indicates the "actual transaction-level deviations identified via controls testing."

Considering only documented evidence, support theory indicates that the packed evidence underlying design deficiencies produces less severe auditor evaluations than the unpacked evidence underlying operating deficiencies – contrary to professional guidance (AICPA 2007 ¶ 62-74). However, documented evidence is not the only evidence auditors can access. Auditors can also access evidence from mentally picturing the deficiency (Dijksterhuis, Macrae, and Haddock 1999). When people have the ability and proclivity to mentally unpack categorical evidence into detailed components, judgment differences disappear between events with packed and unpacked evidence (Van Boven and Epley 2003). The effect of unpacked evidence on judgment is not necessarily exogenously specified. Consequently, we expect that auditors who easily and clearly picture a design deficiency are mentally unpacking it. When a design deficiency is mentally unpacked, we expect design and operating deficiencies are assessed similarly. For auditors who do not easily and clearly picture a design deficiency, we expect support theory's prediction holds: design deficiencies are evaluated as less severe than operating deficiencies. Accordingly, we hypothesize the following.

HYPOTHESIS 1: When auditors can (cannot) easily picture a control deficiency, assessed control risk and increases to substantive tests are no different (lower) for design deficiencies and (than for) operating deficiencies.

Aside from evidence content, there is a second type of evidence that stems from mentally picturing a deficiency: the ease of picturing the deficiency (Schwarz, Bless, Strack, Klumpp, Rittenauer-Schatka and Simons 1991). A significant psychology and marketing literature find that ease of evidence retrieval from memory influences judgment independent of evidence content (Schwarz et al. 1991, Schwarz 1998; Wanke, Bohner, and Jurkowitsch 1997; Tybout, et al. 2005; Danziger, Moran, and Rafaely 2006).⁴ Further, Menon and Raghurir (2003) find that ease of evidence retrieval from memory is an unintentional, subconscious judgment input. Applying these findings to the audit environment, we expect that auditors evaluating a deficiency attempt to mentally picture the deficiency, and we expect that auditors subconsciously account for the ease with which this is accomplished. Because deficiency scenarios are negative events, deficiency ease-of-picturing should increase perceived deficiency severity (Tybout et al. 2005). Consequently, we expect that increases in the ease and clarity of picturing a control deficiency increase the control risk assigned to the deficiency and the perceived need for additional substantive tests. Accordingly, we hypothesize the following.

HYPOTHESIS 2: Increases in perceived deficiency ease-of-picturing increase assessed control risk and adjustment to substantive tests.

In sum, we expect that when auditors easily and clearly picture design deficiencies, they picture transaction-level deviations and thereby unpack the documented design deficiency evidence. Additionally, we expect that auditors account for the ease with which they picture design deficiencies. This produces two judgment inputs from picturing design deficiencies: (1) mentally pictured evidence content and (2) the ease of retrieval from memory (Weber and Johnson 2009). For operating deficiencies, we do not expect that auditors acquire new evidence

⁴ Similar findings exist in the accounting literature (Kadous, Krische, and Sedor 2006).

content from picturing the deficiency, because operating deficiency documentation is complete in describing and enumerating transaction-level deviations. Therefore, we expect that picturing operating deficiencies only produces effects associated with ease of retrieval from memory. Since there are two judgment inputs from picturing a design deficiency versus one from picturing an operating deficiency, we theorize that evaluations of design deficiencies increase in severity with increases in ease-of-picturing at a higher rate than do operating deficiencies. Our expectations are modeled in Figure 1.

[Insert Figure 1 here]

As shown in Figure 1, we expect that design deficiencies are evaluated as less severe than operating deficiencies when ease-of-picturing is low and similar to operating deficiencies when ease-of-picturing is high. Holding the underlying transaction cycle constant, professional guidance indicates that control risk for a design deficiency should be equal to or higher than control risk for an operating deficiency (AICPA 2007 ¶ 62-76). Therefore, our expectations modeled in Figure 1 indicate that control risk for design deficiencies is under-assessed (i.e., not set at maximum). Accordingly, our final hypothesis is as follows.

HYPOTHESIS 3: Assessed control risk for design deficiencies is set at less than the maximum.

3. Research methods and participants

The experimental task requires auditors to read case materials, assess inherent, control, and fraud risk, adjust a substantive audit program for accounts receivable, and answer demographic and debriefing questions. Case materials consist of background information about a manufacturing company, summary financial statements, a narrative description of the revenue transaction processing cycle, and information about the presence or absence of control deficiencies identified in the Sarbanes-Oxley (SOX) 404 audit.

Experimental treatments consisting of no deficiency (control group), design deficiency, and operating deficiency are manipulated via the case materials. Both the design and operating deficiencies are described as significant deficiencies embedded in the credit granting procedures of the company. The design deficiency involves an improperly set system parameter that allows salespeople to violate established credit limits for the entire year. The operating deficiency involves the credit manager becoming ill at mid-year, and when out, credit authority is sporadically executed by the sales manager.⁵ In the control group, case materials indicate that “No material weaknesses were observed. No significant deficiencies were observed.”

The system’s inability to *ever* enforce credit checks represents the design deficiency. Conversely, intermittent breakdowns in separation of duties and prescribed credit check procedures represent the operating deficiency. Credit check controls under the design deficiency are “always down”, whereas credit check controls under the operating deficiency are “sometimes down.” Following professional guidance (AICPA 2007 ¶ 75-76), experimental materials contain control testing results for the operating deficiency, but for the design deficiency, the materials indicate that no control testing has been done.⁶ Appendix A contains treatment manipulations.

Experimental materials were developed through discussions with and review by partners at the participating accounting firm. The substantive audit program follows firm methods and terminology, and the initial time budget matches firm standards for low control risk clients.

⁵ We hold the transaction cycle affected by the control deficiency constant. If sets X and Y contain all possible control deviations, x and y, that could stem from a design or operating deficiency, respectively, in cycle S, then for any ordered pair of deviations (x, y) in cycle S, $x \geq y$ in control risk assessment and increases to substantive tests (AICPA 2007 ¶ 62-76).

⁶ For the operating deficiency, experimental materials support partial reliance on controls over customer credit checks. However, no reliance can be placed on controls over customer credit checks for the design deficiency. Therefore, control risk and substantive procedures related to customer credit checks should be greater for the design deficiency than the operating deficiency.

As a final step, a firm partner specializing in controls testing reviewed the experimental materials and agreed that they were realistic and appropriate for the research question.

Within the ranks of partner, manager, and senior, auditors were randomly assigned to the no deficiency, design deficiency, or operating deficiency treatments. We used two different question orders in the experimental materials and we observed no order effects. Experimental administrators read a script introducing the experiment and distributed envelopes containing the experimental materials and instructions. The experiment was administered during summer training programs at several offices of the participating firm. The firm allotted one hour for completion of the experiment.

Variables

Dependent variables consist of the control risk assessment and the increase to substantive procedures that determine the adequacy of the bad debt provision.⁷ Control risk is measured on an 11-point scale anchored at low risk, medium risk (midpoint) and high risk. Increases to substantive procedures for the adequacy of the bad debt provision is measured on an 11-point scale anchored at none, moderate amount (midpoint) and maximum amount.⁸

Independent variables consist of the experimental treatments and a deficiency ease-of-picturing measure. The experimental treatments are coded as dichotomous variables: control group (no deficiency), design deficiency, and operating deficiency. We measure the ease and

⁷ Auditors assessed inherent risk, control risk, and fraud risk. Auditors also adjusted each accounts receivable program step: analytical procedures, the aged trial balance, confirmation procedures, adequacy of the bad debt provision, and cutoff procedures. All substantive program steps were designed and budgeted for low control risk. Per the experimental materials, determining the adequacy of the bad debt provision involved reviewing the aged trial balance, reviewing confirmation exceptions, and examining credit reports for delinquent accounts. A control deficiency in credit checking should result in increased control risk and increased substantive procedures to determine the adequacy of the bad debt provision. As additional analyses, we examine all risk assessments and substantive test adjustments.

⁸ Auditors were also asked to make actual hour adjustments to the substantive procedures time budget, 22 percent of the auditors in our final sample did not make these adjustments. Therefore, we report the 11-point scale responses. However, analyses of actual hour increases produces results inferentially similar to the 11-point scale responses.

clarity of picturing the control deficiency with two questions: (1) “How easy was it to picture control issues creating financial misstatement?”, and (2) “How clear were the risks with respect to control issues?” The ease and clarity questions were based on a process ambiguity instrument created by French et al. (2006) that identifies ease, clarity, confidence, and satisfaction as measured items. We capture all items on the French et al. (2006) instrument using 11-point scales. A factor analysis indicates that ease and clarity produce a construct distinct from the construct produced by satisfaction and confidence. The average of the ease and clarity responses forms the construct, deficiency ease-of-picturing.⁹

Experimental Checks and Participants

We included two experimental check questions to ensure that the auditors understood the parameters in the experimental materials. The first question determines the level of control deficiency identified in the experimental treatment. Auditors in the control group were expected to choose “no deficiency” and auditors in the design and operating deficiency treatments were expected to choose “significant deficiency.” Our second experimental check question determines the level of control risk assumed in the experimental case’s substantive audit program. Auditors were expected to choose “low control risk.” Failure to respond correctly to either of the checks results in elimination from the sample because being unaware of either the control deficiency or the baseline control risk affects adjustment of the substantive audit program.

Participants consist of audit partners, managers, and seniors attending summer training courses sponsored by their local offices. The firm providing participants is a large, international accounting firm that routinely handles audits of U.S. public companies. We have 134 responses.

⁹ Both principle components factoring with varimax rotation and alpha factoring with oblique rotation are done and produce similar results (i.e., measure loadings greater than 0.60 and non-measure loadings less than 0.30). The oblique rotation indicates that factor correlation is less than 0.20. Internal consistency reliability is 0.81 for the ease-clarity construct and 0.86 for the satisfaction-confidence construct (Bagozzi and Kimmel 1995). We find that satisfaction-confidence did not differ between treatments ($p > 0.10$) and its inclusion as a covariate did not change any models’ inferences. Satisfaction-confidence is therefore unreported. Calculating ease-clarity with factor scores or as a simple average produces results that are inferentially similar.

Within the 134 responses, nine responses are incomplete (seven seniors and two managers/partners) and 38 responses contain experimental check failures. Eleven managers/partners and 27 seniors fail to identify the baseline control risk assumed in the substantive audit program and/or fail to identify the level of deficiency in their treatment.¹⁰

Given the high attrition rate, we analyze demographic differences and hypotheses tests with and without attrition. All experience and self-assessed knowledge metrics are equivalent between auditors in and out of our final sample ($p > 0.10$). Additionally, including the experimental check failures in our statistical analyses produces inferentially similar results to those reported with one exception.¹¹

Table 1 provides detail on the participants. Partners, managers, and seniors make up 9, 30, and 48 of the 87 usable responses. As expected, months of audit experience increases with auditor rank. Partners and managers work on approximately two clients with SOX 404 audits, whereas seniors average about 1.5. While partners and managers have the most clients with significant deficiencies, all experience groups have a relatively small portion of clients with a material weakness. As expected, partners and managers have more planning involvement, extent of testing experience and knowledge of AS 5/2 than seniors.

[Insert Table 1 here]

For seniors, we find no significant differences across treatments ($p > 0.10$) on any reported experience or self-assessment metric shown in Table 1. We find no significant differences between partners' and managers' experience and self-assessments, aside from months of

¹⁰ Adjustment of substantive tests in the presence of an internal control deficiency is a complex task and completing the experimental materials required most of the one-hour allotment. Given the data were collected at local office training sessions, some participants experienced client-related distractions (e.g., phone calls) which caused hurried and partial completion of the experimental materials. By including only auditors that "passed" all experimental checks, we insure a minimum level of attentiveness.

¹¹ We report in the other-findings part of results section that seniors in our final sample assessed inherent risk lower in the design deficiency treatment than the operating deficiency treatment – this result does not hold when adding back all experimental check failures.

experience. For the combination of partners and managers, we find no significant differences across treatments ($p > 0.10$) on any reported experience or self-assessment metric with the exception of clients with significant deficiencies. For partners and managers, the control treatment reports fewer significant deficiency clients than did either the design or operating deficiency treatments (0.79, 1.83, 1.73; $p = 0.073$). Given the small number of partners and the similarity between partners and managers in reported experience with SOX 404 audits, we combine partners and managers for statistical analyses.

4. Results

Table 2 presents auditors' mean responses for control risk assessment, substantive test increase for bad debt provisions, and deficiency ease-of-picturing. As compared to auditors in deficiency treatments, auditors in the no-deficiency control group assess control risk lower, increase substantive tests less, and have more difficulty picturing control issues. Our control group results indicate that auditors in our study are sensitive to control risk differences and increase substantive tests as control risk increases. Additionally, deficiency ease-of-picturing differences between the control group and the deficiency treatments confirm that auditors in the deficiency treatments attempted to picture the deficiency. With regard to the design and operating deficiencies, mean differences exist only for partner and manager control risk assessment (5.54 vs. 7.09; $p = .068$, one-tailed). Following our theory, we next test differences between the design and operating deficiency in a model containing deficiency ease-of-picturing.

[Insert Table 2 here]

Table 3 presents regression results for that model.¹² For partners and managers, we find a

¹² For all regressions reported, Shapiro-Wilk tests indicates normalcy of residuals. White's test rejects heterogeneity of residual variance. Variance inflation factor (VIF) scores indicate high colinearity in the interaction terms. Ease-of-picturing is mean centered to reduce collinearity. Therefore, coefficients are interpreted as the difference between design and operating deficiencies when operating deficiency is at the mean of ease-of-picturing. (Chen, Ender, Mitchell and Wells 2003).

main effect for control risk assessment ($b = -2.407, p = 0.005$, one-tailed) and substantive test increases ($b = -2.063, p = 0.023$, one-tailed) that indicates the design deficiency is evaluated as less severe than the operating deficiency. We do not observe an interaction between the design/operating deficiency and ease-of-picturing for either control risk assessment ($b = 0.429, p = 0.175$, one-tailed) or substantive test increases ($b = 0.589, p = 0.121$, one-tailed).¹³ When deficiency ease-of-picturing is controlled, we find that partners and managers assess control risk lower and increase substantive tests less for the design deficiency than for the operating deficiency.

[Insert Table 3]

Figure 2 graphically illustrates the ease-of-picturing effect on partner and manager evaluations of the design and operating deficiency. As shown, control risk assessment for the design deficiency ($b = 1.159, p = 0.001$, one-tailed) and operating deficiency ($b = 0.730, p = 0.011$, one-tailed) increases in ease-of-picturing; as does substantive test increases for the design deficiency ($b = 1.142, p = 0.003$, one-tailed) and operating deficiency ($b = 0.552, p = 0.058$, one-tailed). Partner and manager deficiency evaluations are strongly influenced by the ease and clarity with which a deficiency is mentally pictured. Further, ease-of-picturing influences the magnitude of difference between design and operating deficiency evaluations as shown in Table 4 panel A (Aiken and West 1993). We find that partners and managers, with perceived ease-of-picturing below nine (eight) on a scale ranging from zero-to-ten, assess control risk (increase substantive tests) statistically significantly lower (less) for the design deficiency than for the

¹³ The interaction coefficients indicate ease-of-picturing slope differences between the design deficiency and the operating deficiency.

operating deficiency.¹⁴

[Insert Table 4 and Figure 2 here]

In support of hypotheses 1, we find that partners and managers who *can* easily picture a deficiency assess control risk and increase substantive tests no differently for the design and operating deficiency. But, partners and managers who *cannot* easily picture a deficiency assess control risk higher and increase substantive tests less for the design deficiency than for the operating deficiency. In support of hypotheses 2, we find that partner and manager evaluations increase in severity as perceived ease-of-picturing a deficiency increases. Finally, our results indicate that the confidence interval around partner and manager design deficiency control risk assessments (3.05 – 8.02, $p = 0.99$) does not include the scale maximum (i.e., ten). In support of hypothesis 3, we find that partners and managers do not, on average, assess control risk for the design deficiency at the maximum.

With regard to senior auditors, Table 3 reports no main effect differences between the design and operating deficiency for control risk assessment ($b = 0.229$, $p = 0.402$, one-tailed) or substantive test increases ($b = -0.206$, $p = 0.400$, one-tailed). We do not find an interaction between the design/operating deficiency and ease-of-picturing for control risk assessment ($b = -0.320$, $p = 0.281$, one-tailed), but we do for substantive test increases ($b = -0.930$, $p = 0.031$, one-tailed). In sum, we do not find differences in control risk assessments between the design and operating deficiency for senior auditors, regardless of deficiency ease-of-picturing. However, seniors' substantive test increases do differ between the design and operating deficiency dependent on the level of deficiency ease-of-picturing.

¹⁴ Considering only partners and managers who perceive ease-of-picturing the most difficult (i.e., the 63% who rated ease-of-picturing below 8 on the 11-point scale), we find that evaluations for the design deficiency are lower than the operating deficiency for control risk assessment (3.86 vs. 6.75, $t = 2.474$, $p = 0.014$, one-tailed) and for substantive test increase (2.14 vs. 5.38, $t = 2.338$, $p = 0.018$, one-tailed).

Figure 3 graphically illustrates the ease-of-picturing effect on senior auditor evaluations of the design and operating deficiency. As shown, control risk assessment for the design deficiency ($b = 0.133, p = 0.378$, one-tailed) does not increase in ease-of-picturing, but it does for the operating deficiency ($b = 0.454, p = 0.099$, one-tailed). With regard to substantive test increases, it also does not increase in ease-of-picturing for the design deficiency ($b = 0.088, p = 0.408$, one-tailed), but it does for the operating deficiency ($b = 1.019, p = 0.001$, one-tailed).¹⁵ We do not observe significant differences between the design and operating deficiency for control risk assessment at any level of ease-of-picturing (see Table 4 panel B). However, when ease-of-picturing is high (low) substantive test increases for the design deficiency is lower (higher) than that for the operating deficiency (see Table 4 panel B).

[Insert Figure 3 here]

For senior auditors, control risk is assessed similarly for the design and operating deficiency, regardless of deficiency ease-of-picturing. Senior auditor substantive test increases differs between the design and operating deficiency, but not as predicted in hypothesis 1. Seniors who *can* (*cannot*) easily picture the deficiency increase substantive tests higher (lower) for the design deficiency than for the operating deficiency. Our senior auditor results do not support hypothesis 1. With regard to hypothesis 2, we find that senior auditor evaluations of the operating deficiency increase as deficiency ease-of-picturing increases. However, we find no such relation for design deficiency evaluations. As such, our senior auditor results offer only partial support for hypothesis 2. Finally, our results indicate that the confidence interval around senior auditors' design deficiency control risk assessment ($5.34 - 8.91, p = 0.99$) does not include the scale

¹⁵ For senior auditors, we observe similar effect patterns for deficiency ease-of-picturing on both control risk assessment and substantive test increase. But as noted, we only find a significant interaction between design/operating deficiency and ease-of-picturing for substantive test increase.

maximum (i.e., ten). In support of hypothesis 3, we find that seniors do not on average assess design deficiency control risk at the maximum.

Other findings

Aside from control risk assessment and adjustment to substantive procedures for bad debt provisions, auditors assess inherent risk and fraud risk, and they adjust substantive tests associated with analytical procedures, the aged trial balance, confirmation procedures, and cutoff procedures. We perform regressions with models identical to those in Table 3 on each of these risk assessments and substantive test adjustments. For partners and managers, none of the models are statistically significant ($p > 0.10$). However, for senior auditors we find significant differences between the design and operating deficiency for inherent risk assessment and substantive test increases for receivable confirmations.

As shown in Table 5 panel A, seniors assess inherent risk lower for the design deficiency than for the operating deficiency (4.94 vs. 6.41, $p = 0.071$, two-tailed). Additionally, seniors increase receivable confirmations less for the design deficiency than for the operating deficiency (1.44 vs. 3.65, $p = 0.005$, two-tailed). Consistent with these results, regressions in Table 5 panel B report significant main effects for inherent risk assessment ($b = -1.755$, $p = 0.034$, two-tailed) and receivable confirmations ($b = -1.986$, $p = 0.006$, two-tailed) controlling for ease-of-picturing effects. We do not find an interaction between the design/operating deficiency and ease-of-picturing for inherent risk assessment ($b = -0.584$, $p = 0.229$, two-tailed), but we do for increases to receivable confirmations ($b = -0.686$, $p = 0.100$, two-tailed). When ease-of-picturing is low, increases to receivable confirmations are no different for the design and operating deficiency, but

when ease-of-picturing is high, increases to receivable confirmations are less for the design deficiency than for the operating deficiency.¹⁶

These findings indicate two items of interest. First, a control deficiency in the credit granting process for trade receivables should primarily affect control risk assessment and substantive procedures for the bad debt provision. Partners and managers evaluate the design and operating deficiency differently for control risk assessment and increases to bad debt provisions. On the other hand, senior auditors evaluate the design and operating deficiency differently not only for increases to bad debt provisions, but also for inherent risk assessment and increases to receivable confirmations.¹⁷ Second, we find that most seniors evaluate the design deficiency as less severe than the operating deficiency – similar to most partners and managers – but seniors make these evaluations on risks and substantive procedures disconnected from the deficiency. Our findings suggest that seniors, like partners and managers, are susceptible to support theory bias, but seniors have difficulty determining the source and consequence of a control deficiency.

5. Discussion

AS 2/5 changed financial statement audits by mandating that internal control analyses from the SOX 404 audit be integrated into the financial statement audit (PCAOB 2007 ¶ 6-8, B6). We inform the integration of the internal control and financial statement audits when auditors encounter design and operating deficiencies. Our results indicate that senior auditors under-assess control risk and under-adjust substantive tests for design deficiencies. Further, our results indicate that partners and managers, who are assigned to review senior work, also under-assess

¹⁶ With regard to ease-of-picturing, the inherent risk assessment for the design deficiency ($b = 0.723$, $p = 0.057$, two-tailed) and the operating deficiency ($b = 0.139$, $p = 0.648$, two-tailed) does not increase in ease-of-picturing. Confirmation procedure adjustment also does not increase in ease-of-picturing for the design deficiency ($b = 0.286$, $p = 0.370$, two-tailed), but it does for the operating deficiency ($b = 0.972$, $p = 0.005$, two-tailed).

¹⁷ Given treatments differed only on the internal control deficiency, observed differences in inherent risk as opposed to control risk imply that seniors mixed the two risks, which is consistent with prior literature (Haskins and Dirsmith 1995; Messier and Austen 2000).

control risk and under-adjust substantive tests for design deficiencies. In sum, our results suggest that under-evaluation of design deficiencies can flow through an audit.

A comparison of auditor evaluations of design and operating deficiencies informs the reason auditors under-assess the consequence of design deficiencies. We observe that most auditors evaluate an operating deficiency as more consequential than a design deficiency, contrary to professional guidance (AICPA 2007 ¶ 62-76). A mistake that is theoretically consistent with support theory bias: operating deficiencies' documented control deviations produce more severe evaluations than design deficiencies' documented design flaws.

Our results also report a finding that is new to the control analysis literature. We find an independent effect on auditors' deficiency evaluations for the ease and clarity with which a deficiency is mentally pictured. For partners and managers, we find a strong ease-of-picturing effect across both design and operating deficiencies. Partners and managers who can easily picture a deficiency evaluate the consequence of that deficiency more severely than partners and managers who cannot easily picture a deficiency. This finding is consistent with psychology literature that indicates memory ease-of-retrieval effects are most pronounced in experts (Caruso 2008). With expertise, memory ease-of-retrieval becomes more diagnostic (Caruso 2008, Hansen and Wanke 2008). Conversely, we only find a deficiency ease-of-picturing effect in the evaluation of operating deficiencies for senior auditors. This finding is consistent with seniors having significant experience performing control tests on operating deficiencies and therefore being able to mentally picture operating deficiencies. Our results suggest that seniors do not have enough experience dealing with design deficiencies to mentally picture them and, therefore, they do not consider design deficiency ease-of-picturing effects.¹⁸ Our ease-of-picturing results

¹⁸ Our results are consistent with the literature on the effects of memory ease-of-retrieval (Tybout et al. 2005). For example, Tybout et al. find that students account for memory ease-of-retrieval when asked to picture themselves in a well-known brand of car, but do not when asked to picture themselves in a lesser known brand.

indicate that when an auditor has experience in deficiency evaluation, the ease with which a deficiency is mentally pictured is influential to deficiency evaluations. This finding appears worthy of further research.

Our research has limitations. Audit planning materials are rich, but they are necessarily restricted in this study due to limits on access to the experimental participants and potential maturity effects in our experiments. Our participants came from one firm. Therefore, our results are potentially specific to the firm that provided the participants. Also, we experienced a high participant attrition rate due to experimental check failure and incomplete responses. However, analysis of the excluded auditors indicated no demographic differences with those left in the sample and adding excluded auditors to those in the sample did not change our key findings. Finally, audits usually involve an audit team, and the ability to consult team members can affect audit judgments. In this experiment, we use individual judgments that do not capture dynamic team interactions, but our focus on individual judgments is consistent with prior research in audit judgment and decision making.

SOX requirements have significantly expanded audits of publicly traded firms, and auditors are required to integrate the internal control and financial statement audits (PCAOB 2007 ¶ 6-8, B6). Like Hogan and Wilkins (2008), our results are consistent with control risk prompting changes to the extent of substantive tests in the post-SOX era. However, our results also suggest that auditors struggle when integrating control deficiencies into the financial statement audit depending on the nature of the deficiency. Our results inform practice by indicating significant potential for the underestimation of risk surrounding a design deficiency, and they inform auditors as to their tendencies when confronted with design deficiencies. These are matters that firms should consider addressing in training. By demonstrating how differences in types of control deficiencies affect planning the financial statement audit, we provide a more complete

understanding of the issues underlying the integration of the internal control and financial statement audits.

References

- Aiken, L. S., and S. G. West. 1993. *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks, CA: Sage.
- American Institute of Certified Public Accountants (AICPA). 1983. *Statement on Auditing Standards No. 47: Audit risk and materiality in conducting an audit*. Durham, NC: AICPA.
- American Institute of Certified Public Accountants (AICPA). 2007. *AU Section 319 - Consideration of Internal Control in a Financial Statement Audit*. Durham, NC: AICPA.
- Bagozzi, R., and S. Kimmel. 1995. A comparison of leading theories for the prediction of goal-directed behavior. *British Journal of Social Psychology* 34:437-461.
- Brody, R. G., J. M. Coulter, and A. Daneshfar. 2003. Auditor probability judgments: Discounting unspecified possibilities. *Theory and Decision* 54: 85-104.
- Caruso, E. M. 2008. Use of experienced retrieval ease in self and social judgments. *Journal of Experimental Social Psychology* 44: 148-155.
- Chen, X., P. B. Ender, M. Mitchell, and C. Wells. 2003. *Regression with STATA*. UCLA: Academic Technology Services, Statistical Consulting Group: www.ats.ucla.edu.
- Danziger, S. S. Moran, and V. Rafaely. 2006. The influence of ease of retrieval on judgment as a function of attention to subjective experience. *Journal of Consumer Psychology* 106 (2): 191-195.
- Dijksterhuis, A., C. N. Macrae, and G. Haddock. 1999. When recollective experiences matter: Subjective ease of retrieval and stereotyping. *Personality and Social Psychology Bulletin* (25) 760–768.
- Fischhoff, B., P. Slovic, and S. Lichtenstein. 1978. Fault trees: Sensitivity of estimates failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance* 4: 330–344.
- Fox, C. R., and R. Birke. 2002. Forecasting trial outcomes: Lawyers assign higher probabilities to possibilities that are described in greater detail. *Law and Human Behavior*, 26 (2): 159-173.
- French, D. P., D. Hevey, S. Sutton, A. Kinmonth, and T. M. Marteau. 2006. Comparison information about health risk. *Journal of Health Psychology* 11:497-510.
- Hansen, J. and M. Wanke. 2008. It's the difference that counts: Expectancy/experience discrepancy moderates the use of ease of retrieval in attitude judgments. *Social Cognition* 26 (4) 447-468.

- Haskins, M. E., and M. W. Dirsmith. 1995. Control and inherent risk assessments in client engagements: An examination of interdependencies. *Journal of Accounting and Public Policy* 14:63-83.
- Hogan, C. E., and M. S. Wilkins. 2008. Evidence on the audit risk model: Do auditors increase audit fees in the presence of internal control deficiencies? *Contemporary Accounting Research* 25:219-242.
- Johnson, E. J., J. Hershey, J. Meszaros, and H. Kunreuther. 1993 Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty* 7: 35–51.
- Kadous, K., S. D. Krische, and L. M. Sedor. 2006. Using counter-explanation to limit analysts' forecast optimism. *The Accounting Review* 81 (2): 377-397.
- Kruger, J., and M. Evans. 2004. If you don't want to be late, enumerate: Unpacking reduces the planning fallacy. *Journal of Experimental Social Psychology* 40: 586–598.
- Macchi, L., D. Osherson, and D. H. Krantz. 1999. A note on superadditive probability judgment. *Psychological Review* 106: 210–214.
- Menon, G., and P. Raghurir. 2003. Ease-of-retrieval as an automatic input in judgments: A mere-accessibility framework? *Journal of Consumer Research* 30: 230-243.
- Messier, W. F., and L. A. Austen. 2000. Inherent risk and control risk assessments: Evidence on the effect of pervasive and specific risk factors. *Auditing: A Journal of Practice and Theory* 19:119-131.
- Public Company Auditing Oversight Board (PCAOB). 2007. *Auditing Standard Number 5 - An Audit of Internal Control Over Financial Reporting that is Integrated with an Audit of Financial Statements*. Washington D.C.: PCAOB.
- Redelmeier, D. A., D. J. Koehler, V. Liberman, and A. Tversky. 1995. Probability judgments in medicine: Discounting unspecified possibilities. *Medical Decision Making* 15: 227–230.
- Rottenstreich, Y., and A. Tversky. 1997. Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review* 104 (2) 406-415.
- Russo, J. E., and K. J. Kolzow. 1994. Where is the fault in fault trees? *Journal of Experimental Psychology: Human Perception and Performance* 20: 17–32.
- Savitsky, K., L. Van Boven, N. Epley, W. M. Wight. 2005. The unpacking effect in allocations of responsibility for group tasks. *Journal of Experimental Social Psychology* 41: 447-457.
- Schwarz, N. 1998. Accessible content and accessibility experiences: The interplay of declarative and experiential information judgment. *Personality & Social Psychology Review* 2(2): 87-99.

- Schwarz, N., H. Bless, F. Strack, G. Klumpp, H. Rittenauer-Schatka, and A. Simons. 1991. Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology* 61: 195–202.
- Schwarz, N., and L. A. Vaughn. 2002. The availability heuristic revisited: Ease of recall and content as distinct sources of information. In *Heuristics and biases: The psychology of intuitive judgment*, ed. T. Gilovich, D. Griffin, and D. Kahneman, 103–119. Cambridge: Cambridge University Press.
- Tenney, E. R., H. M. D., Cleary, B. A. Spellman. 2009. Unpacking the doubt in “Beyond a reasonable doubt”: Plausible alternative stories increase not guilty verdicts. *Basic and Applied Social Psychology* 31: 1-8.
- Tversky A., and D. J. Koehler. 1994. Support theory: A nonextensional representation of subjective probability. *Psychological Review* 101 (4): 547-567.
- Tybout, A. M., B. Sternthal, P. Malaviya, G. A. Bakamitsos, and S. Park. 2005. Information accessibility as a moderator of judgments: The role of content versus retrieval ease. *Journal of Consumer Research* 32: 76-85.
- Wanke, M., G. Bohner, and A. Jurkowitsch. 1997. There are many reasons to drive a BMW: Does imagined ease of argument generation influence attitudes? *Journal of Consumer Research* 24: 170–77.
- Van Boven, L., and N. Epley. 2003. The unpacking effect in evaluative judgments: When the whole is less than the sum of the parts. *Journal of Experimental Social Psychology* 39: 263-269.
- Weber E. U., and E. J. Johnson. 2009. Mindful judgment and decision making. *Annual Review of Psychology* 60: 53-85.

TABLE 1
Participant profile: mean (std. dev.)

Panel A: Auditor experience

	Partners	Managers	Seniors
Number of participants	9	30	48
Months of auditor experience	219.44 (76.42)	93.47 (53.44)	34.69 (15.70)
Number of clients with SOX 404 audits	2.11 (1.54)	2.10 (2.46)	1.46 (1.75)
Number of clients with significant deficiencies	1.78 (1.79)	1.29 (1.12)	0.84 (1.31)
Number of clients with material weaknesses	0.33 (0.50)	0.36 (0.56)	0.28 (0.63)
Number of clients with audit planning involvement	9.22 (2.33)	9.54 (1.07)	6.42 (3.22)
Number of clients on which audit extent decisions were made	9.00 (3.00)	8.52 (2.87)	5.44 (3.45)

Panel B: Auditor self-assessments

	Partners	Managers	Seniors
Knowledge of AS 5/2	7.33 (1.58)	6.41 (1.86)	4.38 (2.45)
Understood case materials			
Control group	9.00 (1.00)	8.50 (1.57)	8.87 (1.46)
Design and implementation deficiency	7.00 (2.65)	9.20 (1.32)	8.88 (1.26)
Operating effectiveness deficiency	9.00 (1.00)	9.50 (0.54)	7.65 (2.87)

Notes:

Experience measures were captured on 11-point scales with endpoints labeled “none” and “10 or more.” Self-assessments were captured on 11-point scales: knowledge of AS2/5 endpoints were labeled “Poor” and “Excellent” and understood case material endpoints were labeled “no” and “yes.” Differences across treatments were tested with ANOVA and nonparametric Kruskal-Wallis tests. For seniors, we found no significant differences across treatments ($p > 0.10$) on any reported experience or self-assessment metric. We found no significant differences between managers and partners experience and self-assessments, aside from months of experience. For the combination of managers and partners, we found no significant differences across treatments ($p > 0.10$) on any reported experience or self-assessment metric with the exception of clients with significant deficiencies. For managers and partners, the control treatment reported fewer significant deficiency clients than did either design deficiency treatment or the operating deficiency treatment (0.79, 1.83, 1.73, respectively) ($p = 0.073$).

TABLE 2
Variable means (std. dev.) and mean tests

Participants	Constructs	Treatment Means (std. dev.)			ANOVA (<i>F</i> -value)	DD vs. OD (<i>t</i> -value)
		Control	DD	OD		
All Auditors	Control Risk	3.10 (2.412)	6.41 (2.732)	7.04 (2.411)	20.657*	0.930
	Substantive Test Increase	1.37 (2.157)	4.45 (2.798)	5.07 (2.581)	18.177*	0.933
	Ease-of- Picturing	5.43 (1.888)	7.16 (1.632)	6.77 (1.960)	7.194*	0.798
	n	30	29	28		
Partners/Managers	Control Risk	3.20 (2.210)	5.54 (2.933)	7.09 (2.212)	8.181*	1.531‡
	Substantive Test Increase	1.53 (2.386)	3.92 (2.842)	5.18 (2.639)	6.627*	1.175
	Ease-of- Picturing	5.13 (1.913)	7.23 (1.787)	6.41 (2.177)	4.121†	1.028
	n	15	13	11		
Seniors	Control Risk	3.00 (2.673)	7.13 (2.419)	7.00 (2.598)	12.946*	0.140
	Substantive Test Increase	1.20 (1.971)	4.88 (2.778)	5.00 (2.622)	11.592*	0.144
	Ease-of- Picturing	5.73 (1.879)	7.09 (1.551)	7.00 (1.837)	2.874†	0.153
	n	15	16	17		

Notes:

Control = No control deficiency; DD = Design deficiency; OD = Operating deficiency.

Control risk is measured on an 11-point scale with endpoints and midpoints labeled “low risk,” “medium risk,” and “high risk.” Substantive test increase is measured from a low control risk baseline on 11-point scales with end and midpoints labeled “none,” “moderate amount,” and “maximum amount.” Ease-of-Picturing was measured as the mean of two 11-point scales: (1) ease-of-picturing control issues with endpoints “not easy” and “very easy,” and (2) clarity of the risks associated with control issues with endpoints “not clear” and “very clear.”

* Significant at the 0.01 level (one-tailed).

† Significant at the 0.05 level (one-tailed).

‡ Significant at the 0.10 level (one-tailed).

TABLE 3
 Regressions on control risk assessments and substantive test increases for the bad debt provision:
 Coefficients (*t*-statistic) reported

	All Auditors		Partners/Managers		Seniors	
	Control Risk	Substantive Test	Control Risk	Substantive Test	Control Risk	Substantive Test
Intercept	6.835* (11.727)	4.629* (8.154)	7.256* (11.480)	5.307* (7.452)	6.835* (10.842)	4.629* (8.314)
Design Deficiency (DD)	0.229 (0.271)	0.206 (0.249)	-2.407* (2.737)	-2.063 [†] (2.082)	0.229 (0.250)	0.206 (0.254)
Control Group	-3.397* (3.797)	-2.994* (3.437)	-4.214* (4.492)	-3.790* (3.586)	-3.397* (3.501)	-2.994* (3.504)
Ease-of-Picturing	0.454 [‡] (1.416)	1.019* (3.264)	0.730 [‡] (2.411)	0.552 [‡] (1.620)	0.454 [‡] (1.309)	1.019* (3.328)
DD x Ease-of-Picturing	-0.320 (0.633)	-0.930 [†] (1.887)	0.429 (0.948)	0.589 (1.155)	-0.320 (0.585)	-0.930 [†] (1.924)
Control x Ease-of-Picturing	0.032 (0.069)	-0.537 (1.190)	-0.835 [†] (1.988)	-0.563 (1.191)	0.032 (0.063)	-0.537 (1.213)
Partner/Manager	0.421 (0.457)	0.678 (0.755)				
DD x Partner/Manager	-2.637 [†] (2.019)	-2.269 [†] (1.784)				
Control x Partner/Manager	-0.818 (0.590)	-0.796 (0.589)				
Ease-of-Picturing x Partner/Manager	0.276 (0.589)	-0.466 (1.022)				
DD x Ease-of-Picturing x Partner/Manager	0.749 (1.042)	1.520 [†] (2.168)				
Control x Ease-of-Picturing x Partner/Manager	-0.867 (1.307)	-0.026 (0.041)				
n	87	87	39	39	48	48
F-statistic	6.267	6.412	8.173	5.626	5.971	8.388
(<i>p</i> -value)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Adjusted R ²	0.403	0.409	0.486	0.378	0.346	0.440

Notes:

Control = No control deficiency; DD = Design deficiency; OD = Operating deficiency.

The DD and the control group are coded as dummy variables and OD is held out of each regression.

Therefore, DD and the control group coefficients represent least square mean differences with OD.

Partner/manager is coded as follows: partner/manger = 1, senior auditor = 0. Ease-of-picturing is mean centered to reduce collinearity. Variable construction is defined in Table 2.

* Significant at the 0.01 level (one-tailed).

† Significant at the 0.05 level (one-tailed).

‡ Significant at the 0.10 level (one-tailed).

TABLE 4

Differences between operating and design deficiencies at levels of ease-of-picturing for control risk assessment and substantive test increases for the bad debt provision

Panel A: Partners/Managers

Ease-of-Picturing	Control Risk		Substantive Test Increase	
	Difference (DD – OD)	<i>t</i> -value (DD – OD = 0)	Difference (DD – OD)	<i>t</i> -value (DD – OD = 0)
0 (Low)	-5.254	1.63 [‡]	-5.974	1.64 [‡]
1	-4.825	1.73 [†]	-5.384	1.71 [†]
2	-4.396	1.86 [†]	-4.780	1.80 [†]
3	-3.937	2.03 [†]	-4.206	1.91 [†]
4	-3.538	2.26 [†]	-3.616	2.05 [†]
5 (Medium)	-3.189	2.56 [*]	-3.027	2.21 [†]
6	-2.680	2.80 [*]	-2.437	2.26 [†]
7	-2.251	2.57 [*]	-1.849	1.87 [†]
8	-1.822	1.80 [†]	-1.259	1.10
9	-1.393	1.07	-0.670	0.46
10 (High)	-0.964	0.58	-0.801	0.04

Panel B: Senior Auditors

Ease-of-Picturing	Control Risk		Substantive Test Increase	
	Difference (DD – OD)	<i>t</i> -value (DD – OD = 0)	Difference (DD – OD)	<i>t</i> -value (DD – OD = 0)
0 (Low)	2.356	0.59	6.378	1.82 [†]
1	2.035	0.59	5.448	1.80 [†]
2	1.745	0.59	4.518	1.76 [†]
3	1.494	0.58	3.588	1.70 [†]
4	1.074	0.57	2.657	1.59 [‡]
5 (Medium)	0.753	0.53	1.727	1.36 [‡]
6	0.433	0.41	0.797	0.85
7	0.113	0.13	-0.133	0.17
8	-0.208	0.20	-1.063	1.17
9	-0.528	0.38	-1.994	1.63 [‡]
10 (High)	-0.849	0.46	-2.924	1.80 [†]

Notes:

Control = No control deficiency; DD = Design deficiency; OD = Operating deficiency.

Differences and their related *t*-values are determined from the regression equations reported in Table 3.

* Significant at the 0.01 level (one-tailed).

† Significant at the 0.05 level (one-tailed).

‡ Significant at the 0.10 level (one-tailed)

TABLE 5
Senior auditor inherent risk assessments and substantive tests increase for receivable confirmations

Panel A: Mean (std. dev.) and mean tests

	Treatment Means (Std. Dev.)			ANOVA (<i>F</i> -value)	DD vs. OD (<i>t</i> -value)
	Control	DD	OD		
Inherent Risk	5.00 (2.360)	4.94 (2.235)	6.41 (2.265)	2.194 [‡]	1.852 [†]
Receivable Confirmations	1.47 (1.642)	1.44 (1.825)	3.65 (2.714)	5.793*	2.968*
n	15	16	17		

Panel B: Regressions analyses: Coefficients (*t*-statistic) reported

	Inherent Risk	Receivable Confirmations
Intercept	6.361* (11.561)	3.293* (7.003)
Design Deficiency (DD)	-1.755 [†] (2.193)	-1.986* (2.905)
Control Group	-1.005 (1.190)	-1.792 [†] (2.483)
Ease-of-Picturing	0.139 (0.459)	0.972* (3.762)
DD x Ease-of-Picturing	0.584 (1.222)	-0.686 [‡] (1.680)
Control x Ease-of-Picturing	-0.256 (0.584)	-0.934 [†] (2.50)
n	48	48
<i>F</i> -statistic	2.045	5.934
(<i>p</i> -value)	(0.092)	(0.000)
Adjusted R ²	0.100	0.344

Notes:

Control = No control deficiency; DD = Design deficiency; OD = Operating deficiency

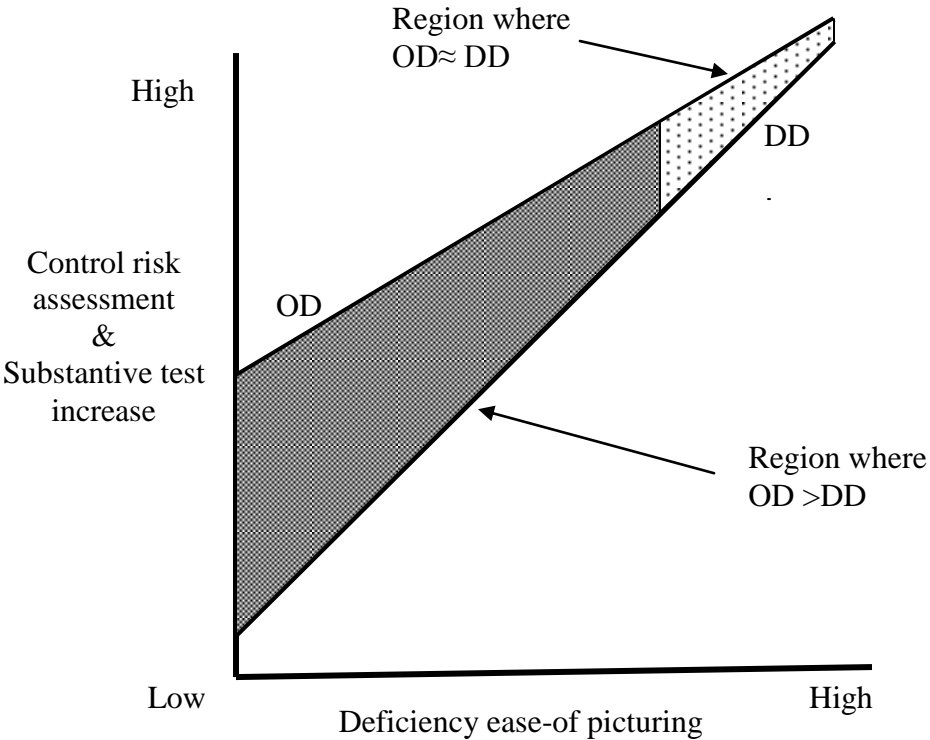
The DD and control are coded as dummy variables and OD is held out of each regression. Therefore, DD and the control coefficients represent least square mean differences with OD. Ease-of-picturing is mean centered to reduce collinearity. Variable construction is defined in Table 2.

* Significant at the 0.01 level (two-tailed).

[†] Significant at the 0.05 level (two-tailed).

[‡] Significant at the 0.10 level (two-tailed)

Figure 1 Expected relation between evaluations of design and operating deficiencies and deficiency ease-of-picturing

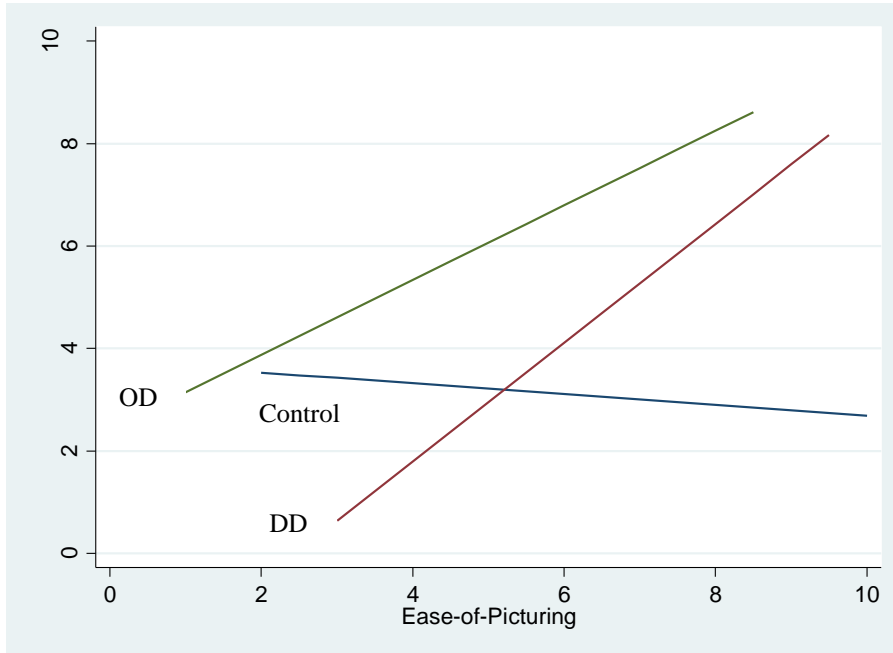


Notes:

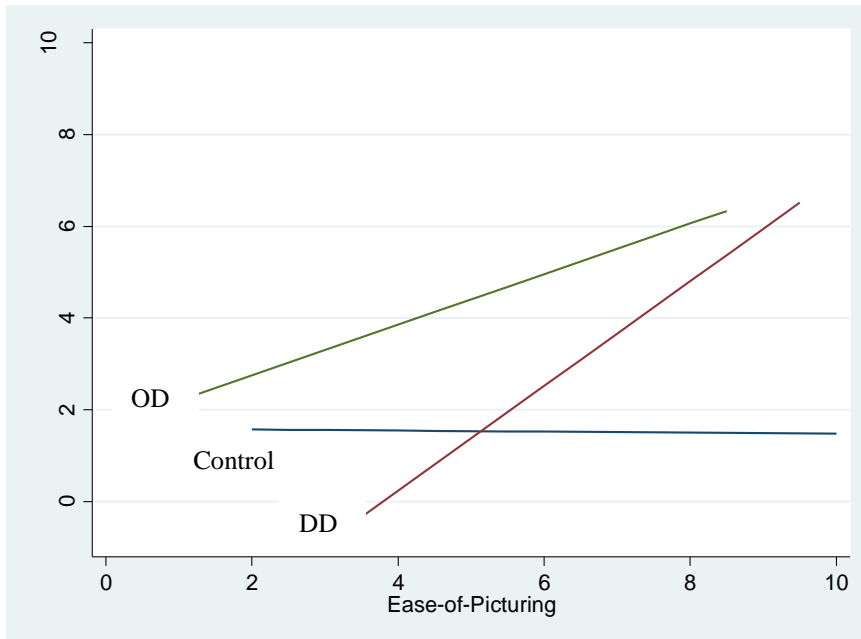
Control = No control deficiency; DD = Design deficiency; OD = Operating deficiency.

Figure 2 Partners and managers deficiency evaluations by deficiency ease-of-picturing

Panel A: Control Risk Assessment



Panel B: Substantive test increase for the bad debt provision



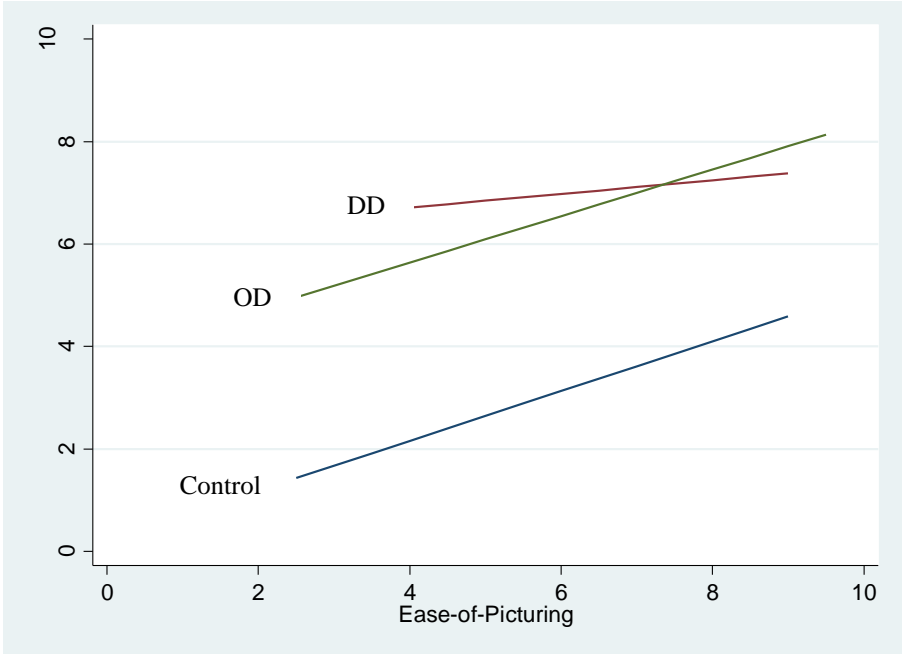
Notes:

Control = No control deficiency; DD = Design deficiency; OD = Operating deficiency.

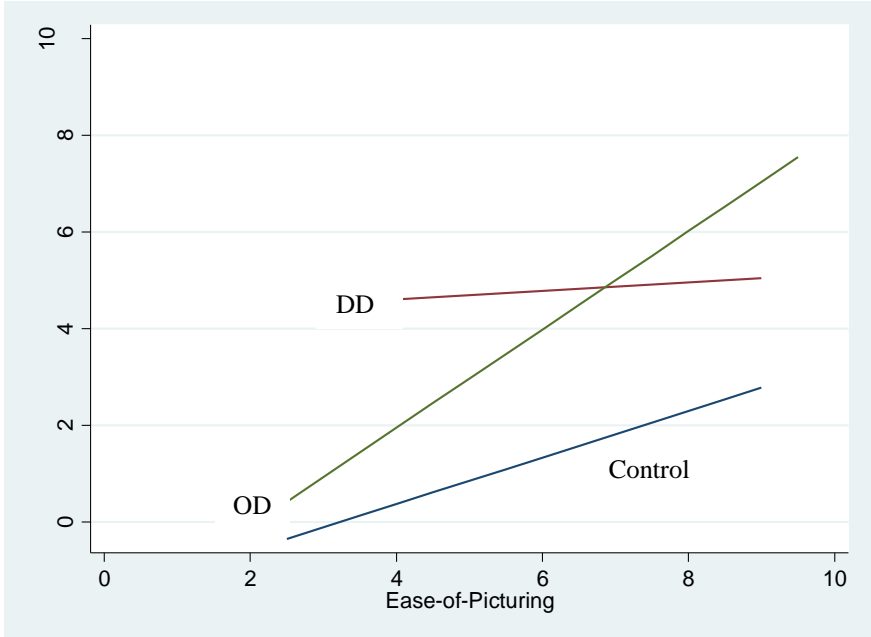
Graphs based on the regressions reported in Table 3.

Figure 3 Senior auditors deficiency evaluations by deficiency ease-of-picturing

Panel A: Control Risk Assessment



Panel B: Substantive test increase for the bad debt provision



Notes:

Control = No control deficiency; DD = Design deficiency; OD = Operating deficiency.

Graphs based on the regressions reported in Table 3.

Appendix
Experimental Materials (Manipulations)

Control Deficiencies. No material weaknesses were observed. No significant deficiencies were observed. **[control treatment]** However a significant deficiency was identified in the customer credit approval process and was reported to management and the board of directors. The details of the significant deficiency are outlined in the following report. **[OD & DD treatments]**

CONTROL DEFICIENCY EVALUATION FORM [OD & DD TRMTS ONLY]

This Form assists the auditor or issuer in documenting and classifying individual process/transaction-level control deficiency.

Business Cycle:	Sales
Control Objective:	Specific approval of credit for all new customers and ensuring customers stay below their approved limit to decrease risk of bad debts to the company.
Control Category: <i>(Preventative vs. Detective)</i>	Preventative Control
Deficiency Description:	[OD] The credit manager was out due to illness on a number of occasions and all customer credit items awaiting his approval were given an override in the system to allow processing and fulfillment of orders. [DD] The accounting software notifies sales personnel when customers will exceed their credit limit; however, these personnel can simply click a box that says “proceed” and process the order. This allows customers to exceed their credit limit. The system was not set up to stop these orders from being processed.
Risk Description:	Customers [OD] exceeded [DD] can exceed their credit limit, increasing company exposure to uncollectible accounts and bad debts.
Deficiency Type: <i>(Design & Implementation vs. Operating Effectiveness)</i>	Operating Effectiveness. Company has policies and procedures in place, but they were not operating as designed. <ul style="list-style-type: none"> • Testing indicated that credit limits were exceeded on 27 customers during the year due to overrides. • Testing indicated that three (3) new customers were shipped their orders before undergoing credit approval. Design and Implementation. Company has <u>not</u> correctly implemented the control within the accounting system, although it is possible to implement this control.
Compensating Controls	The controller monitors overdue customer account balances.
Financial Statement Line Item affected:	<ul style="list-style-type: none"> • Reserve for Doubtful Accounts • Bad Debt Expense
Likelihood of Misstatement: <i>(Reasonable Possibility)</i>	We determined that there is a <u>reasonable possibility</u> that this control deficiency could cause financial reporting misstatements.
Potential Magnitude of Misstatement: <i>(Significant Misstatement and Material Misstatement)</i>	We determined that if misstatements occur, the amount could be a <u>significant misstatement</u> , but not a material misstatement. Materiality is set at 5% of net income which is approximately \$80,000.
Deficiency Classification: <i>(Deficiency, Significant Deficiency, and Material Weakness)</i>	Significant Deficiency
Remediation Plan / Mgmt Response:	[OD] Management is aware of this issue and has agreed to assign a trained back-up employee who can perform credit issuance controls in the credit manager’s absence. [DD] Management is aware of this issue and offered to change the defaults within the accounting system.
Follow-Up Procedure & Finding:	[OD] We performed a review at year-end and confirmed that credit control overrides occurred at least one additional time since notifying management. The controls have not been operating effectively as of the 12/31/2007 reporting date. [DD] We reviewed the system defaults, noting that the system has NOT been re-set to disallow orders to be processed over the credit limit.
Deficiency Status: <i>(Open vs. Resolved)</i>	Open.

When did we observe / periods affected by deficiency:	January –December of 2007.
---	----------------------------

Additional information from the BAS manager responsible for IT control analyses on the engagement. We reviewed documentation of the procedures over recording of revenue and the related receivables. **[OD]** We noticed that there was a problem in operating effectiveness due ineffective backup procedures and personnel. The credit manager was diagnosed with lung cancer in July of 2007 and immediately began treatment which caused him to miss a lot of work. When he is out, they never know if it is going to be for just a day or for much longer. The company was getting a backlog of orders waiting for credit approval. In an effort to keep customers satisfied, CAP had to be able to release the orders. So, the system administrator has given the sales manager access to the credit system to release these orders as necessary. The issue is that the sales manager is busy with his regular duties so he often releases the orders without doing the credit check. He does do them if he has time, and the credit manager tries to catch up on them when he gets back to work, but this is weeks or months after the initial release of goods and by that time additional sales have usually been made.

The client was made aware of the operating effectiveness flaw. However, due to staffing issues and the credit manager’s continuing illness, the situation persisted throughout the reporting period. Management has indicated that an assistant credit manager will be hired and trained in 2008 to rectify the situation.

[DD] When looking at the design of controls in the system, we noticed that there was a discrepancy in what the procedures said was being accomplished and what was actually occurring within the system. The default within the client’s system is for a warning message to appear when an order would exceed the credit limit for a particular customer. However, this is only a warning, and clicking on the option to “proceed” will place the order. Basically, sales personnel can override every single credit overage without any oversight. This was an obvious design/implementation flaw, so we did not test the operating effectiveness of the control.

The client was made aware of the design/implementation flaw. However, due to IT staffing and the need for a consultant to make system changes, the production program was not changed by year-end. Management has indicated that the needed system changes will be made in 2008.