The Outcome Effect and Professional Skepticism

Joseph F. Brazel
North Carolina State University

SUMMARY: Despite the importance placed on professional skepticism by the accounting profession and regulators, the failure of auditors to exercise an appropriate level of skepticism continues to be a global issue. This article summarizes a recent study by Brazel, Jackson, Schaefer, and Stewart (2016) that examines a potential barrier to skepticism: that outcome knowledge biases supervisors' evaluations of skeptical behavior. Holding a staff member's skeptical judgments and acts constant, Brazel et al. (2016) find that superiors on engagement teams evaluate the staff's skeptical behavior based on whether the staff's investigation of an issue ultimately identifies a misstatement. The evidence suggests that evaluators penalize auditors who employ an appropriate level of skepticism, but do not identify a misstatement. Collectively, Brazel et al. (2016) depict an evaluation system that may inadvertently discourage skepticism amongst auditors in the field.

Keywords: audit; evaluation; hindsight bias; outcome effect; professional skepticism.

I. INTRODUCTION

This article summarizes a recent study titled “The Outcome Effect and Professional Skepticism” (Brazel et al. 2016). Specifically, I review the study's motivation and hypotheses, research method, and results. I then discuss several of the study's more significant implications. Most importantly, the study depicts an evaluation system that may be inadvertently discouraging professional skepticism amongst auditors in the field.

I thank two anonymous reviewers for their valuable comments.

Editor's note: Accepted by Lisa Milici Gaynor.

Submitted: August 2018
Accepted: November 2018
Published Online: December 2018
II. MOTIVATION AND HYPOTHESES

Professional Skepticism

Professional skepticism requires the auditor to maintain a questioning mind and critically assess audit evidence throughout the planning and performance of an audit (PCAOB 2006; IAASB 2004). As auditors exercise more professional skepticism, they may require more evidence to justify their audit opinions (Nelson 2009). Auditors should “not be satisfied with less-than persuasive audit evidence” (IAASB 2004), and superiors on the engagement team should ensure that their subordinates exercise an appropriate level of skepticism.

Despite the recognized importance of professional skepticism by the accounting profession and regulators, auditors’ failure to exercise a sufficient level of skepticism continues to be a global issue (PCAOB 2008; AIU 2010; ASIC 2010; EC 2010; IAASB 2012; PCAOB 2012a). A lack of skepticism has been cited as a cause of audit failures, SEC and PCAOB enforcement actions, and malpractice claims filed against auditors (e.g., Carmichael and Craig 1996; Beasley, Carcello, and Hermanson 2001; Anderson and Wolfe 2002; Messier, Kozloski, and Kochetova-Kozloski 2010; PCAOB 2012b). While most parties agree that efforts to improve professional skepticism are needed (e.g., KPMG 2012; PCAOB 2012a; PwC 2012), the underlying causes and potential solutions for insufficient skepticism remain unclear.

The underlying cause(s) of insufficient skepticism must first be understood in order to effectively identify how to increase professional skepticism in auditors. Most of the accounting literature focuses on auditor-specific traits (e.g., knowledge, innate characteristics) that may result in insufficient skepticism (e.g., Hurtt, Eining, and Plumlee 2011). Prior research has yet to fully identify how firms’ evaluation systems may be either encouraging or discouraging skeptical behavior (IAASB 2008; Hurtt et al. 2013).1

An Unexplored Barrier to Professional Skepticism—Outcome Effect Bias in Auditor Evaluations

“Outcome effects” refer to situations where the knowledge of outcomes influences evaluators’ judgments in the direction of the outcome (Tan and Lipe 1997). Although outcome knowledge can be diagnostic, incorporating it into evaluations of decision quality effectively biases the evaluation when the outcome is not diagnostic of the quality of decision making (Mertins, Salbador, and Long 2013). Given a costly event (e.g., medical matters, failure to meet targets, insolvency), evaluators exhibiting outcome bias will focus on outcomes, rather than the uncertainty inherent in a decision at the time that decision was made (e.g., Baron and Hershey 1988; Emby, Gelardi, and Lowe 2002; Ghosh and Lusch 2000).2

1 One reason for this lack of research may be the difficulty associated with measuring and observing both evaluation systems and skeptical behavior.

2 Hindsight bias is a similar phenomenon to outcome effects, as noted by Lipe (1993) and Mertins et al. (2013). However, although both result from the possession of outcome knowledge and can be examined jointly (e.g., Baron and Hershey 1988; Emby et al. 2002), the mechanism through which the outcome knowledge affects evaluations of decision quality differs. The direct impact of outcome knowledge on evaluations is referred to as outcome effects, whereas the effect of outcome knowledge on the judged probability of outcomes (and, indirectly, evaluations) is referred to as hindsight bias. Brazel et al. (2016) confirm that their results are attributable to outcome effects (versus hindsight bias).
To understand how outcome effects in auditor evaluations may negatively impact skepticism, we must consider the associated costs of skeptical behavior. The obvious costs associated with *not being skeptical* and failing to identify a misstatement make it seem intuitive that heightened skepticism is always more beneficial than less skepticism. For example, the failure to identify a material misstatement in the financial statements may result in a restatement. In turn, the audit firm may lose the client, the auditor may lose their job, and investors may bring lawsuits against the firm (Nelson 2009).

The costs associated with *being skeptical*, on the other hand, are less obvious. Highly skeptical auditors may decrease the risk to the firm by reducing the risk that material misstatements will go undetected, but their skepticism may result in additional inquiries or procedures and budget overruns that the audit firm may not be able to recover. Additionally, performing unplanned or atypical audit procedures (compared to prior year audits) may anger or frustrate the client, as they are required to respond to unanticipated inquiries and evidence requests. While either of these costs may lead to consequences for the firm (e.g., less profitability and/or a strained relationship with the client), they may also lead to direct consequences for the staff auditor.

Professional skepticism is a behavior that, although encouraged by the profession, does not always produce the same outcome (e.g., sometimes it leads to the identification of a misstatement and other times it may not). Consider a situation where an auditor observes a red flag or inconsistency when assessing audit evidence and exercises an appropriate level of skepticism by performing additional testing. Conducting an investigation would be consistent with exercising professional skepticism; however, it requires added effort from the auditor and the client and does not ensure that a misstatement will be found. It is possible that the investigation leads to an acceptable explanation for an unusual pattern of facts observed, such that no audit adjustment is necessary. In short, the auditor has incurred the costs associated with skepticism (went over budget and/or strained client relations), but has not experienced the “benefit” of identifying a misstatement.

When assessing an auditor’s decision making (e.g., evaluating their skeptical judgments and actions), the appropriate criterion for evaluating the auditor is not what they found, but what the evidence suggested they *might* find (Lipshtiz and Barak 1995; Mertins et al. 2013). Still, research on outcome effects suggests that when costs are incurred, an auditor’s performance evaluation may be influenced more by the outcome of their skeptical behavior (i.e., whether or not a misstatement is found) than by whether they engaged in the appropriate level of skeptical behavior (i.e., appropriately identified and investigated a red flag or inconsistency). As such, skeptical judgments and acts that incur costs, but identify no misstatement, may be viewed less favorably by superiors. On the other hand, skeptical behavior that results in the detection of a misstatement will likely be viewed more favorably by superiors as there is a tangible benefit from the investigation. That is to say, when a misstatement is found, the end justifies the means. Hypothesis 1 is as follows:

**H1:** Superiors will evaluate skeptical auditors more negatively (*positively*) when they do not (*do*) identify a misstatement.

**Auditor Consultation—Removing the Bias of Outcome Effects from the Evaluation Process**

If H1 is supported and superiors’ evaluations of skeptical auditors are biased by outcome effects, then potential solutions to debias evaluations are a relevant consideration. Prior research
finds that evaluators start with outcome knowledge and work backward to connect the causal links that led to the outcome—increasing the salience of outcome-consistent information (Mertins et al. 2013). As mentioned above, when there is a perceived benefit as a result of incurring costs, H1 predicts that evaluators will reward the decision maker for incurring the costs. However, when there is no benefit to match with the costs, the decision maker may be penalized as cognitive reconstruction increases the perceived inevitability of the outcome, making it seem as though the decision maker should have “known it all along.”

Glover and Prawitt (2014) suggest that creating a “culture of consultation” may enhance the appropriate application of skepticism in the field. Accordingly, one potential solution to debias superiors’ evaluations of skeptical auditors would be to involve the evaluator in the decision-making process in order to reduce the perceived inevitability of the outcome should the outcome be negative. Auditors can do this through informal consultations with their superior before engaging in the skeptical action. Hypothesis 2 is as follows:

**H2:** When subordinate auditors consult with their superiors during the course of exercising skepticism, the outcome effect in auditor evaluations is reduced.

### III. METHOD

#### Participants

The study’s hypotheses were tested in an experiment where the participants consisted of 96 audit seniors from an international accounting firm. Brazel et al. (2016) (hereafter, “we” or “our”) administered the experiment while participants attended a training session. On average, our participants had 30 months of audit experience and had conducted three performance evaluations of staff auditors under their supervision.4

#### Description of Experimental Context

The experimental materials placed participants in the position of the lead senior on the hypothetical audit engagement of Madison, Inc., a publicly traded manufacturing company with multiple divisions. Participants learned that Madison had been an audit client for ten years and had received an unqualified audit opinion each year. The materials also provided information about the auditor-client relationship, including: (1) Madison was a large audit client in the office, (2) the budget was very tight and there was pressure to keep fees down, (3) there had been few historical audit adjustments, and (4) Madison expected the audit to run smoothly and asked for explanations when the nature, timing, and extent of audit procedures changed.

Participants were informed that their task was to evaluate the performance of a third-year staff member, Sam, who worked under their supervision. Sam’s responsibilities on the audit included, among other tasks, performing substantive analytical procedures related to the revenue account

---

3 The literature on professional skepticism makes a distinction between the skeptical judgment (i.e., identifying inconsistencies) and the skeptical action (i.e., investigating inconsistencies) (Nelson 2009). We are suggesting that consultation between the skeptical judgment and the skeptical action may reduce outcome effects in the evaluation.

4 See Brazel, Gimbar, Maksymov, and Schaefer (2018) for a replication of Brazel et al. (2016) with a sample of more experienced audit seniors.
for the Sporting Goods division of Madison. In the past, the analytical procedures had incorporated prior year Madison financial information and industry financial trends to develop an expectation for the division’s revenues.

The literature related to professional skepticism stresses the importance of evaluating inconsistent evidence (e.g., IAASB 2004). In our experimental setting, the information sources used in prior year testing (Madison’s own past financial performance and industry financial data) were consistent with the revenue account reported by the division in the current year. Non-financial measures (NFMs) for Madison, such as number of employees and square footage of facilities, were not considered in prior years. In all conditions, participants were informed that Sam incorporated NFMs into his analytical procedures in the current year and noticed an inconsistency between the revenue account and related NFMs. Thus, Sam’s decision to collect and consider alternative, inconsistent evidence by itself can be considered skeptical behavior.

Sam then chose to investigate the inconsistency. As such, we hold the staff auditor’s skeptical judgment (issue identification) and skeptical action (additional investigation) constant between all experimental conditions. The investigation of the inconsistency caused Sam to go over budget and strain relations with management (i.e., incur the costs of skepticism [Nelson 2009]) in all conditions.

**Manipulated Variables**

The first variable manipulated between participants is whether Sam’s investigation uncovered a misstatement (this variable is referred to as **OUTCOME**). In the no misstatement condition, participants were told the following:

Sam found that the inconsistency described above was a result of the Sporting Goods division outsourcing some operations overseas. Sam made several inquiries into the matter and collected additional audit evidence, which eventually led to a conclusion that there were no misstatements in this revenue account.

In the misstatement condition, participants were told the following:

Sam found that the inconsistency described above was a result of the Sporting Goods division outsourcing some operations overseas. Sam made several inquiries into the matter and collected additional audit evidence, which eventually led to a conclusion that a significant misstatement existed in this revenue account as revenues were being recognized prematurely at the overseas operation.

This variable reflects a common feature of the real-world audit environment—the exercise of professional skepticism may or may not result in the identification of a misstatement.

Knowing this, auditors may attempt to mitigate an adverse evaluation by consulting with their supervisor prior to investigating evidence inconsistencies. Investigating inconsistencies takes time for the auditor and client, so some auditors may consult with their supervisor before doing so. As such, the second independent variable manipulates consultation between participants at three levels (this variable is referred to as **CONSULT**). In the no consultation condition, participants were told the following:

5 To be clear, participants were not asked to perform analytical procedures themselves. Instead, they were asked to evaluate the performance of an audit staff member who had performed those procedures.
Without consulting with you, Sam chose to investigate the inconsistency between the growth in Sporting Goods revenues and the decreases in the number of employees and production space instead of relying on other sources like the industry trends that support the reported revenue growth, which is what was done in prior years.

In the moderate consultation condition, participants were told the following:

Sam consulted with you about investigating the inconsistency between the growth in Sporting Goods revenues and the decreases in the number of employees and the square footage of production facilities. Your response to Sam’s consultation about investigating the inconsistency was “use your professional judgment.” Sam chose to investigate the inconsistency between the growth in revenues and the decreases in the number of employees and production space instead of relying on other sources like the industry trends that support the reported revenue growth, which is what was done in prior years.

In the high consultation condition, participants were told the following:

Sam consulted with you about investigating the inconsistency between the growth in Sporting Goods revenues and the decreases in the number of employees and the square footage of production facilities. Your response to Sam’s consultation about investigating the inconsistency was “I approve of you investigating the inconsistency.” After receiving your approval, Sam investigated the inconsistency between the growth in revenues and the decreases in the number of employees and production space instead of relying on other sources like the industry trends that support the reported revenue growth, which is what was done in prior years.

We manipulated consultation at three levels because each level reflects a different option open to auditors in practice. The no consultation condition reflects the option to rely solely on one’s own judgment rather than seek guidance. The moderate consultation condition reflects the option to inform their supervisor about a situation and allow the supervisor to provide guidance if needed (this amounts to keeping the supervisor “in the loop”). The high consultation condition reflects the option to inform their supervisor about a situation and get their approval before incurring the costs associated with investigating the inconsistency.

**Dependent Variable**

Our dependent measure (EVAL) is participants’ evaluations of Sam, which was elicited using the following question:

Based on the information presented on the prior pages, how would you evaluate Sam’s overall performance?

Participants responded on an 11-point response scale ranging from −5 to +5, with the left endpoint labeled “Below Expectations,” the right endpoint labeled “Above Expectations,” and the midpoint labeled “Met Expectations.”

**IV. RESULTS**

Figure 1 illustrates the mean responses for EVAL in all six experimental conditions. Visual inspection of Figure 1 suggests that OUTCOME has a large influence on participants’ performance evaluations, consistent with H1. Contrary to H2, Figure 1 reveals that, when no misstatement is
The dependent variable is participants’ overall performance evaluation, which is their response to the question “How would you evaluate Sam's overall performance?” (responses are provided on an 11-point scale ranging from −5 to +5 with the left endpoint labeled “below expectations”, the right endpoint labeled “above expectations”, and the midpoint labeled “met expectations”). The manipulated variables are defined as follows: OUTCOME (manipulated between participants as either [1] Sam found that there was no misstatement or [2] Sam found that there was a significant misstatement) and CONSULT (manipulated between participants at three levels as either [1] Sam investigated the inconsistency without consulting with the senior, which is referred to as no consultation, [2] Sam investigated the inconsistency after consulting with the senior and being told to use professional judgment, which is referred to as moderate consultation, or [3] Sam investigated the inconsistency after consulting with the senior and receiving approval, which is referred to as high consultation). See the Method Section (Section III) for a full description of the experiment.

found, Sam is evaluated less favorably at every level of consultation (0.35 versus 2.61 in the no consultation condition, 1.69 versus 3.18 in the moderate consultation condition, and 1.19 versus 2.67 in the high consultation condition). Even when Sam kept his evaluator “in the loop” or received approval for the investigation, he was still penalized if he did not identify a misstatement. Importantly, in all conditions we observed that evaluators indicated a strong and equal belief that the inconsistency should have been investigated by Sam. Thus, even though evaluators are

6 Discussions with representatives at the participating firm revealed an important cultural factor related to performance evaluations at the firm. The firm’s performance evaluations are positively skewed. In practice, the majority of auditors are evaluated as performing “Above Expectations.” Auditors are evaluated as “Met Expectations” far less often, and only a small number are evaluated as falling “Below Expectations.” To confirm this anecdotal information, as part of a separate study, we asked 62 audit seniors where they would rank in their peer group if they consistently received “Met Expectations” ratings. Of the 62 audit seniors, approximately 76 percent indicated that they would be in the bottom half of their peer group. Given this cultural factor, evaluations indicating that Sam has met expectations should be viewed as an indication of relatively poor performance.

7 Statistical analyses (i.e., ANOVA results) documented in Brazel et al. (2016) confirm that the evidence supports H1 and does not support H2.
charged with ensuring their subordinates exercise an appropriate level of skepticism and, in this case, agreed that the matter under consideration warranted additional investigation, the staff auditors that found no misstatement still received lower evaluations.

---

**Additional Analyses**

In Brazel et al. (2016), the primary tests of hypotheses are accompanied by a host of additional analyses that I will briefly summarize here. First, we demonstrate the process by which **OUTCOME** affects **EVAL**. When the outcome of professional skepticism is *not* the identification of misstatement, evaluators perceive that the investigation provided less benefit to the audit team. In turn, a lower perceived benefit causes the evaluator to frame the cost of the investigation as “lost time” (versus a “normal cost” of the audit). Last, the framing of skepticism that does not yield a misstatement as “lost time” causes the superior to provide a lower performance evaluation of Sam.

Second, we performed a similar experiment with corporate managers responding to skeptical audit staff that did or did not identify a misstatement in their company’s financial statements. The results of the experiment provide additional evidence that a skeptical auditor is penalized (management conveys a less favorable evaluation to the partner on the engagement) when their skeptical behavior does not identify a misstatement.

Third, we executed an additional experiment with a separate sample of auditors that manipulated the level of audit committee support rather than consultation. The expectation was that higher (versus lower) audit committee support would insulate the auditor from the costs of skepticism (budget overruns and strained management relations) and alleviate the outcome effect. Similar to consultation, we found that the audit committee support manipulation did not reduce the outcome effect on evaluations.

---

**V. IMPLICATIONS**

Our study should be of interest globally to accounting firms, regulators, standard setters, and the investing community as they attempt to find ways to increase the professional skepticism exercised by auditors in the field. Although the profession calls for more skepticism, the underlying culture may inhibit such behavior if auditors are punished for being skeptical when it turns out there is no misstatement. We find that performance evaluation systems punish auditors for exhibiting appropriate levels of skepticism by changing the frame in which the costs of skepticism are viewed by evaluators. In addition, although one might think that consultation could provide an easy fix for outcome bias in evaluations, we find that even when subordinates consult with their supervisor and receive permission to proceed, supervisors are unable to purge outcome bias from their evaluations of audit staff.

Our results indicate that outcome effects will result in evaluators penalizing skeptical behavior when no misstatement is found. The immediacy of this negative consequence may affect staff behavior in the face of uncertain outcomes. An auditor in the field may perceive that they would be better off not being skeptical when the likelihood of detecting a misstatement is uncertain. Glover and Prawitt (2014) pose this as a possible threat to professional skepticism. This reasoning also suggests that audit firms’ evaluation and reward systems may be inadvertently discouraging skepticism amongst auditors in the field. In order for outcome effects in the evaluation process to affect the level of professional skepticism in the field, audit staff must *anticipate* that their behavior will be evaluated differently depending on the outcome. Such anticipation may cause auditors in the field to avoid investigating red flags/inconsistencies unless they are obvious or the identification of a misstatement is a “sure thing.”
We tested the extent to which outcome effects are anticipated by audit staff with a survey of Master of Accounting (MACC) students from two large public universities. Of these participants, 93 had no audit work experience and 43 had audit work experience (usually through an internship). The MACC participants who had audit experience had worked in public accounting for an average of six months and they had all received at least one performance evaluation. Results of the survey, described in more detail in Brazel et al. (2016), confirmed that participants anticipated that their performance evaluations would be significantly lower when their skepticism did not yield a misstatement (versus when it identified a misstatement). Such anticipation was found amongst both inexperienced and experienced MACC participants, which suggests that auditors understand from the onset of their careers that skeptical behavior will be evaluated with outcome knowledge in mind.

Given these concerns, in a separate study Brazel, Gimbar, Maksymov, and Schaefer (2018) first replicate Brazel et al. (2016) and find that evaluators once again reward audit staff who exercise appropriate levels of skepticism and identify a misstatement (positive outcome). However, when no misstatement is identified (negative outcome), evaluators penalize staff who exercise appropriate levels of skepticism. One factor causing this outcome effect may be that exercising skepticism typically causes budget overages due to additional testing. Hence, Brazel et al. (2018) also examine whether formally attributing the budget overage to skeptical judgments and actions in the audit budget file reduces outcome effects. However, Brazel et al. (2018) are unable to reduce the outcome effect by documenting the staff’s skepticism alongside their budget overage. Thus, it is clear that the outcome effect in the context of evaluating skepticism is very robust. Consequently, an internal, firm-level training that makes evaluators aware of this bias (and its potential consequences) may be more effective than a subordinate-driven solution (e.g., increased consultations, documenting skepticism in the budget). Last, our findings should inform human resource policies that shape the evaluation processes employed by audit firms.

The findings of Brazel et al. (2016) should spur future research in relation to the evaluation of auditor skepticism. For example, when audit partners are the evaluators and audit managers are the skeptical subordinates, are outcome effects less pronounced or do they persist to more senior members of the engagement team? Also, will the disclosure of critical audit matters in the audit report, which undoubtedly require enhanced skepticism, cause evaluators to be more apt to view the costs of skepticism as a normal cost of the audit versus lost time?

REFERENCES


