THE ACCOUNTING REVIEW
Vol. 85, No. 4
2010
pp. 1273–1301

American Accounting Association
DOI: 10.2308/accr.2010.85.4.1273

Auditors' Use of Brainstorming in the Consideration of Fraud: Reports from the Field

Joseph F. Brazel
North Carolina State University

Tina D. Carpenter
The University of Georgia

J. Gregory Jenkins
Virginia Polytechnic Institute and State University

ABSTRACT: Audit standards require auditors to conduct fraud brainstorming sessions on every audit. The Public Company Accounting Oversight Board has raised concerns about auditors' fraud judgments and the quality of their brainstorming sessions. We develop a measure of brainstorming quality to examine how it affects auditors' fraud decision-making processes. We test our measure using field survey data of auditors' actual brainstorming sessions for 179 audit engagements. Respondents report considerable variation in the quality of brainstorming in practice. We find some evidence that high-quality brainstorming improves the relations between fraud risk factors and fraud risk assessments. We also determine that brainstorming quality positively moderates the relations between fraud risk assessments and fraud-related testing. Our results suggest that the benefits of brainstorming do not apply uniformly, because low-quality sessions likely incur the costs of such interactions without receiving the attendant benefits. By documenting best practices from high-quality brainstorming sessions, our findings can inform auditors on how to improve their consideration of fraud.

Keywords: *brainstorming quality; fraud; fraud risk assessments; fraud risk response*.

We owe special thanks to Steve Kachelmeier (editor) and two anonymous reviewers. We are grateful for the helpful comments provided by Chris Agoglia, Ben Ayers, Ann Backof, Linda Bamber, Michael Bamber, Mark Beasley, Jodi Bellovary, Grey Burnett, Rex Yuxing Du, Brooke Elliott, Jenny Gaver, Steve Glover, Rick Hatfield, Rich Houston, Kathy Hurtt, Karla Johnstone, Bill Messier, Mitch Oler, Tom Omer, Linda Parsons, Mark Peecher, Denis Pelletier, Robin Pennington, Dave Piercey, Velina Popova, Bob Ramsay, Jane Reimers, Steve Salterio, Jaime Schmidt, Marjorie Shelley, Chad Simon, Jason Smith, CJ Song, Edward Swanson, Chris Wolfe, William Wright, Mark Zimbelman, and workshop participants at the University of Illinois, Texas A&M University, University of Nevada, Las Vegas, Baylor University, North Carolina State University, The University of Alabama, the 2007 AAA Annual Meeting, the North Carolina Office of the State Controller, First Citizens Bank, and the 2008 Auditing Section Midyear Meeting. We also thank Deloitte & Touche, Ernst & Young, Grant Thornton, KPMG, and PricewaterhouseCoopers for their support and participation in this project.

Editor's note: Accepted by Steven Kachelmeier.

Submitted: July 2008
Accepted: December 2009
Published Online: June 2010

I. INTRODUCTION

Fraud detection is among the highest priorities for the accounting profession and standard-setters (Elliott 2002; PCAOB 2003, 2007, 2008). Statement on Auditing Standards (SAS) No. 99 (AICPA 2002), *Consideration of Fraud in a Financial Statement Audit,* requires brainstorming sessions on every audit as a means of improving auditors' fraud judgments. However, the Public Company Accounting Oversight Board (PCAOB) has expressed concerns about the actual conduct and quality of brainstorming sessions in practice (PCAOB 2007). Because these sessions are a relatively new requirement, little is known about how they occur in practice and how they influence fraud-related judgments. Accordingly, our study employs a field survey of actual audit engagements to pursue three research objectives. First, we study how audit teams are conducting brainstorming sessions in practice and develop a measure of brainstorming quality. This descriptive analysis should provide a basis for further modeling of brainstorming and future empirical research in team decision-making. Second, we investigate the comprehensive fraud decision-making process that audit teams are employing in practice, from the consideration of client fraud risk factors to their development of fraud risk assessments and fraud-related audit testing. Third, we examine how the quality of these brainstorming sessions influences auditors' consideration of fraud and whether variation in this quality has important consequences for the resulting fraud judgments.

This investigation is important for several reasons. First, prior psychology research suggests that the effectiveness of judgments and decisions likely depends on the quality of a team's interaction (e.g., Stasser 1999). We leverage the survey method to collect complex and rich data about brainstorming session quality and its underlying characteristics. Our measure of brainstorming session quality is multifaceted (i.e., 21-item scale) and is developed from a review of the psychology literature on brainstorming and team decision-making, the accounting literature on auditors' judgments and fraud, recent PCAOB inspection reports, and professional guidance. Our brainstorming session-quality measure consists of three overall elements: attendance and communication, structure and timing, and engagement team effort. To our knowledge, this is the first study in accounting or elsewhere to develop a measure of brainstorming quality and empirically examine its effects on team judgments.

Second, prior experimental studies have examined certain parts of the brainstorming process, but have not examined the entire decision-making process (e.g., Carpenter 2007; Hoffman and Zimbelman 2009). Moreover, the accounting literature suggests that the benefits derived from brainstorming depend on the brainstorming method used (Carpenter 2007; Hoffman and Zimbelman 2009; Lynch et al. 2009; Hunton and Gold 2010). Carpenter (2007) and Hoffman and Zimbelman (2009) document benefits to brainstorming versus not brainstorming. However, contemporary studies by Lynch et al. (2009) and Hunton and Gold (2010) suggest that the type of brainstorming is an important consideration for the benefits to be realized, as they find that some effective brainstorming methods are not predominant in practice.[1] While this prior research has

[1] Lynch et al. (2009) conclude that brainstorming effectiveness is higher for teams using group support systems relative to teams using traditional face-to-face brainstorming. In our sample, 80 percent of our observations used traditional face-to-face brainstorming. Hunton and Gold (2010) find that both the nominal group (where there is *no* communication among team members) and round robin brainstorming techniques are more effective than open brainstorming. However, in our sample, we find that 91 percent of our engagements employed open brainstorming. Thus, while these two studies provide means in which brainstorming can be improved, our descriptive data suggest that such changes would be significant and likely a substantial challenge for firms to undertake. Our finding of a preponderance of face-to-face communication and open brainstorming in practice likely suggests that unidentified institutional factors cause audit teams to employ these two potentially less effective methods (e.g., open brainstorming [versus round robin] allows a partner to better control the discussion).

provided some evidence on auditor fraud judgments in experimental settings, there is no empirical evidence on how audit teams are implementing brainstorming sessions in practice and whether the relations among their team judgments are improved by the *quality* of brainstorming sessions required by SAS No. 99. We contribute to the literature by using a field survey to collect auditors' client-specific brainstorming session data and the related fraud judgments. Because this methodology allows us to collect the underlying dynamics of audit teams' actual brainstorming sessions, and the resulting risk factors identified, fraud risk assessments, and related testing, we are able to examine multiple fraud decisions and relationships across a number of diverse audit engagements. Thus, we provide insight into the comprehensive fraud decision-making process that audit teams employ in practice.

Third, our investigation is important to auditors and standard-setters. PCAOB inspection reports indicate variation among audit teams' brainstorming sessions, and auditors' testing is often not responsive to their fraud risk assessments (PCAOB 2007). Our study empirically assesses the PCAOB's findings by documenting how auditors are implementing SAS No. 99 in practice and by investigating the consequences of brainstorming quality on the link between fraud risk assessments and subsequent testing. Thus, this study's findings can inform audit practice and standard-setters as they reconsider the professional guidance on auditors' consideration of fraud (PCAOB 2004, 2007).

Our theoretical model is based on the psychology and accounting literatures, PCAOB inspection reports, and professional guidance. Based on this model, we hypothesize that brainstorming quality positively moderates the relations between (1) fraud risk factors and related fraud risk assessments and (2) fraud risk assessments and subsequent audit testing. We test our hypotheses with data from 179 audit engagements for which highly experienced audit partners and managers from the Big 4 firms and an international firm completed a field survey.

Prior experimental research has typically found that when fraud risk factors are present, fraud risk assessments tend to be higher (e.g., Nieschwietz et al. 2000). Further, prior experimental research has documented that auditors have difficulty responding to fraud risk assessments by altering the nature of their testing (e.g., Zimbelman 1997; Glover et al. 2003). Actual audit engagements are rich in context and the fraud risk assessment is typically determined after engagement *team* deliberation. Such an environment is difficult to create in an experimental setting. Therefore, there is a need to complement experimental fraud research with research employing non-experimental methods. Our field survey results indicate that fraud risk factors identified for a diverse sample of clients are positively correlated with auditor fraud risk assessments. In addition, we find some evidence that brainstorming quality positively moderates the relations between fraud risk factors and fraud risk assessments. We also find that auditors' responses to fraud risk assessments are contingent on the quality of brainstorming sessions. Specifically, when audit teams engage in higher quality brainstorming sessions, fraud risk assessments are more positively associated with changes to the nature, staffing, timing, and extent of fraud-related audit procedures.

While several experimental studies have predicted a link between fraud risk and related audit testing, they have typically not found this link for the nature or type of procedures employed (e.g., Zimbelman 1997; Glover et al. 2003; Asare and Wright 2004). We complement this literature by documenting the important role brainstorming quality can play in improving the link between auditors' fraud risk assessments and their subsequent testing; including the choice of the *nature* of procedures. We consequently illustrate that important and tangible benefits can be derived from higher quality brainstorming. On the other hand, our analyses suggest that under-auditing can occur when brainstorming session quality is low. Thus, some engagement teams likely incur the cost of brainstorming without receiving the attendant benefits of such interactions. These new findings triangulate and extend prior studies by offering important insights on the costs and benefits of fraud brainstorming.

Legal restrictions prohibit the PCAOB from distributing descriptive evidence about brainstorming sessions in practice (PCAOB 2007). Thus, it is difficult for individuals outside the PCAOB to discern if the PCAOB's concerns about brainstorming quality are widespread or perhaps isolated to only a few audit engagements. Our study captures brainstorming data that is often not documented and thus not available for PCAOB inspection teams to review (e.g., participation levels of engagement team members). Our survey data validate the concerns raised by the PCAOB by documenting considerable variation in the reported quality of brainstorming sessions in practice. More importantly, we find evidence that there are consequential effects of conducting low- and high-quality sessions. We also analyze data related to the misappropriation of assets and fraudulent financial reporting that were detected by the audit teams in our sample (i.e., instances of actual fraud detection). In additional analyses, we examine audit team effectiveness by analyzing the links between *ex ante* fraud judgments and *ex post* fraud detection. Last, our data reveal best practices for how auditors can improve the quality of their brainstorming sessions (e.g., attendance of an IT audit specialist, occurrence of the session early in the planning process).

Section II describes the background and hypotheses development. Sections III and IV provide the methods and results of the study, respectively. Section V concludes the paper.

II. BACKGROUND AND HYPOTHESES DEVELOPMENT

A Model of Auditors' Use of Brainstorming in the Consideration of Fraud

Based on SAS No. 99, as well as prior accounting and psychology research, we develop a model of auditors' use of brainstorming in the consideration of fraud. SAS No. 99 provides guidance to improve the likelihood that auditors will detect fraud using a multi-phase approach (AICPA 2002, 2003). First, auditors collect information related to the risk of material misstatement due to fraud (e.g., data on equity-based compensation). Using such information, auditors brainstorm to identify fraud risk factors (e.g., incentives), synthesize this information to develop a fraud risk assessment, and develop a response to the risk assessments such as altering the staffing of the engagement, or modifying the nature, timing, and extent of audit procedures. Brainstorming sessions are intended to aid auditors in linking fraud risk factors to risk assessments and, in turn, foster the development of appropriate audit responses. As such, the model depicted in Figure 1 indicates that brainstorming should influence both phases of the fraud decision-making process such that the relations among fraud risk factors, risk assessments, and responses are positively moderated by the quality of the brainstorming session.

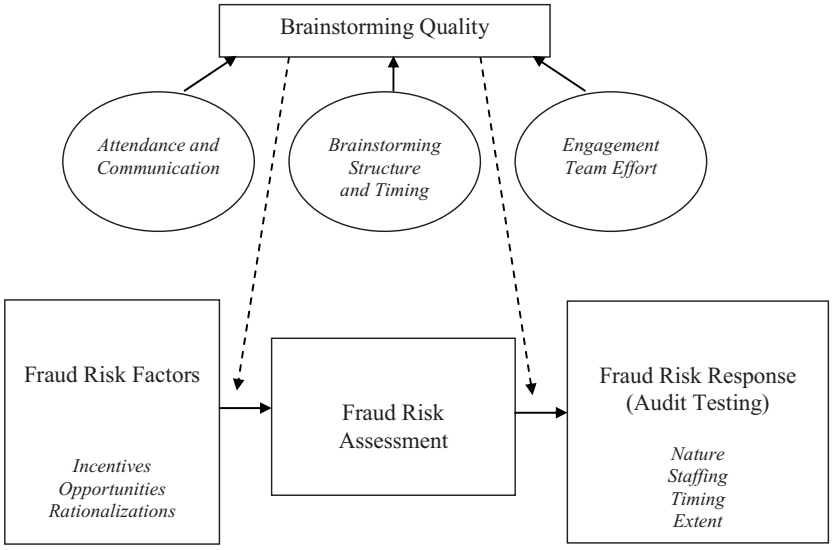
Fraud Risk Factors

Fraud risk factors are commonly categorized along three dimensions of the fraud triangle: incentives, opportunities, and rationalizations (AICPA 2002). Auditors synthesize their understanding of these risk factors with other information (e.g., preliminary analytical procedures) to develop a fraud risk assessment (AICPA 2002). Prior archival research on fraud risk factors documents relations between fraudulent financial reporting and financial statement elements, corporate governance quality, nonfinancial measures, and other observable data (e.g., Beasley 1996; Beneish 1997; McVay 2006; Brazel et al. 2009).

Fraud Risk Assessments

Experimental studies have examined the relationships between fraud risk factors and auditors' fraud risk assessments. While auditors' risk assessments are generally higher when fraud risk factors are present (Nieschwietz et al. 2000), more recent studies report that context has an important influence on auditors' judgments. Wilks and Zimbelman (2004) find that auditors who decompose fraud risk assessments are more sensitive to incentive and opportunity risks than auditors who provide holistic fraud risk assessments, but only in a low fraud risk context. Asare

FIGURE 1
Auditors' Use of Brainstorming in the Consideration of Fraud



This figure provides a model of audit teams' use of brainstorming in their consideration of fraud. This model is based on psychology and accounting research as well as the SAS No. 99 framework.

and Wright (2004) observe that, given a case where fraud was perpetrated, auditors who use a fraud risk checklist make less effective fraud risk assessments than those who use no checklist. Since the context of the experimental setting was important in these studies, and these studies were conducted before the implementation of SAS No. 99, it is important to re-examine the relations between fraud risk factors and fraud risk assessments using non-experimental methods. Further, none of these studies consider the possible influence of brainstorming quality on the links between risk factors and risk assessments.

Fraud Risk Responses

Although SAS No. 99 requires auditors to design audit procedures that are responsive to fraud risk assessments, research has not consistently found such a relationship. For example, Zimbelman (1997) illustrates limited evidence of an increase in the extent of audit procedures and no evidence of a change in the nature of procedures in response to fraud risk assessments. While Glover et al. (2003) find stronger evidence of an increase in the extent of audit procedures; they also fail to observe evidence of a change in the type of procedures employed in response to fraud risk. Additionally, Asare and Wright (2004) report no association between fraud risk assessments and planned audit procedures. However, these studies were conducted prior to SAS No. 99 and, thus, exclude the possibility that brainstorming could improve the sensitivity of fraud risk responses to fraud risk assessments.

Brainstorming Quality

Carpenter (2007) and Hoffman and Zimbelman (2009) experimentally study the influence of brainstorming sessions on the link between fraud risk and fraud risk assessment and the link between risk assessment and the nature of the audit procedures selected, respectively. Carpenter (2007) examines hierarchical teams that brainstorm and concludes that teams generate lists with more quality fraud risks than do individual auditors. Further, these teams provide more effective fraud risk assessments (i.e., higher when fraud was present than when it was not) than auditors working alone. Hoffman and Zimbelman (2009) investigate the role that brainstorming plays with regard to audit managers' planned audit procedures. They document that brainstorming (versus no brainstorming) leads to more effective modification of standard audit procedures in a specific high fraud risk case. While both of these studies provide important insights, neither captures the collective fraud process (i.e., the relations among fraud risk factors, risk assessments, and responses), nor do they examine actual audit teams where partners' judgments are likely very important.[2]

More importantly, contemporary experimental research by Lynch et al. (2009) and Hunton and Gold (2010) find that the brainstorming methods that are predominant in practice (e.g., face-to-face communication and open brainstorming) are less effective than alternative methods (e.g., electronic communication and round robin brainstorming). For a sample of audits where open brainstorming and face-to-face discussion were the norm, we examine if higher brainstorming quality can mitigate the negative effects documented by these contemporary studies. Furthermore, the PCAOB has expressed concern about brainstorming quality in practice (PCAOB 2007). However, because the PCAOB is legally restricted from distributing data from their inspections, it is difficult for those outside the PCAOB to discern if brainstorming quality issues are widespread or isolated, and the extent that session quality varies between engagements. Therefore, it is an open empirical question as to whether differences in brainstorming quality actually found in practice are large enough to lead to meaningful differences in audit teams' fraud risk assessments and audit testing.

There is a substantial body of psychology literature on the role of brainstorming. In general, this research reports mixed findings, with evidence suggesting that brainstorming can lead to both process gains and losses (e.g., Hill 1982; Diehl and Stroebe 1987). Most of these experimental studies were conducted with college students who held no task-specific expertise or knowledge and were asked to brainstorm about mundane problems (e.g., Diehl and Stroebe 1987). However, SAS No. 99 requires auditors with specific training and experience to conduct brainstorming sessions, a task with much at risk. Past studies of brainstorming also did not include hierarchical teams, eliminating the potential costs and benefits of normal hierarchies (cf., Hinsz et al. 1997). Hierarchical teams of auditors are likely to develop synergies by building on one another's ideas (e.g., Carpenter 2007). Conversely, the hierarchy of an audit team might exacerbate problems that can occur during brainstorming, such as evaluation apprehension, group think, and production blocking (e.g., Beasley and Jenkins 2003).

Psychology research has documented links between cooperative outcome interdependence, effectiveness of team communications, and resultant decisions. Under cooperative outcome interdependence, team members succeed or fail together, directly benefiting from each other's performance (De Dreu 2007). Audit teams have cooperative outcome interdependence because their

[2] Neither Carpenter (2007) nor Hoffman and Zimbelman (2009) include audit partners in their brainstorming teams. Given the authority level assigned to audit partners, it is likely that they wield substantial influence over fraud-related judgments. In our sample, partners attended every brainstorming session and were the largest contributors to the sessions (see Table 3). Also, whether the partner led the session and the amount of partner contribution were positively correlated with a number of fraud-related judgments (e.g., nature of testing; see Table 5). Our use of survey data from actual audit engagements allows us to include the role of partners.

team interaction in a brainstorming session depends on cooperation as opposed to competition. Cooperative outcome interdependence is related to the teams' ability to discuss shared and unshared information (Stasser 1999). Thus, the quality of information sharing or brainstorming among team members is likely to influence the effectiveness of the discussion and, in turn, the decisions of the team.

Building on professional guidance and the accounting and psychology literatures, our model in Figure 1 posits that the quality of the brainstorming session (1) positively moderates the relations between auditors' fraud judgments and (2) is based on three elements: attendance and communication, brainstorming structure and timing, and engagement team effort.

Attendance and Communication

Attendance and communication affect the quality of team interaction and discussion. Postmes et al. (2001) suggest that teams that promote critical thought are more likely to ensure careful evaluation of information and arrive at higher quality decisions. Schippers et al. (2003) find that the quality of team communication is enhanced when teams are more diverse and when they encourage open communication. Stasser (1999) illustrates that openness to ideas improves decision quality by encouraging the sharing of information. In line with this research, we expect that as more members of the engagement team attend and participate in the brainstorming session, there will be greater diversity and more sharing of information. This should improve the overall quality of the session and the responsiveness of fraud judgments.

Brainstorming Structure and Timing

The structure and timing of team discussions are important to the quality of team judgments. For example, increased time pressure exacerbates the tendency of teams to seek cognitive closure, reduces motivation, and increases reliance on heuristics (e.g., Kruglanski and Freund 1983; De Dreu 2003). Further, De Dreu (2003) suggests that pressure reduces team motivation to process information in a systematic and deliberate way, leading to reduced team effectiveness. Although time pressure is ever-present in the audit environment, auditors can potentially mitigate its negative effects by, for example, conducting brainstorming sessions early in the planning phase of the audit rather than later. Sessions held earlier in the planning process are expected to positively influence auditors' fraud judgments as the engagement team will have more time to implement the ideas endorsed during the session.

Engagement Team Effort

Engagement team effort is also an important determinant of the quality of teams' brainstorming sessions. There are two qualitatively different mechanisms of information processing: shallow, less critical, heuristic information processing and more argument-based, effortful, systematic information processing (Chaiken and Trope 1999). If the goal is to discover popular beliefs and preferences, team interaction tends to be preference-driven, leaving little opportunity for innovative decision making. In contrast, information-driven teams employ effortful systematic processing, communicating and integrating relevant information, to arrive at higher quality decisions (Stasser and Birchmeier 2003). Such an effortful process focus is evident in SAS No. 99-related guidance, which suggests that, for example, auditors individually identify risks and potential audit responses prior to brainstorming (AICPA 2003). These requirements should enhance each team member's involvement in the fraud audit process, augment their client-specific knowledge, and improve their contribution to the brainstorming session. Therefore, we expect that the more effort auditors exert prior to and during brainstorming, the higher the quality of team interaction and resultant fraud-related judgments.

In summary, higher quality brainstorming sessions are expected to improve (1) the links between fraud risk factors and related assessments and (2) the relations between fraud risk assessments and subsequent testing. That is, we expect session quality to positively moderate both relations, as stated in the following hypotheses:

H1: Fraud risk factors are more positively related to fraud risk assessments in high-quality brainstorming sessions than in low-quality brainstorming sessions.

H2: Fraud risk assessments are more positively related to fraud risk responses in high-quality brainstorming sessions than in low-quality brainstorming sessions.

III. METHOD

Participants

We approached contacts in the Big 4 firms and one other international firm to request participation in the study. Each firm agreed to participate. A secure Internet survey link was sent to each contact who distributed the survey to partners, directors, senior managers, and managers in their offices. A total of 367 auditors received the survey link, of which 179 completed the survey, resulting in a response rate of 48.8 percent. This rate exceeds the response rates in other auditing surveys (e.g., 16 percent response rate by Nelson et al. [2002]; 29.7 percent response rate by Dowling [2009]); however, it is lower than the exceptionally high 70.5 percent achieved by Gibbins et al. (2001).[3] Participants held positions of partner (56), director (2), senior manager (60), and manager (61). All 179 brainstorming sessions in our study reported to have an engagement partner in attendance.

Research Instrument

Similar to the method used by Gibbins et al. (2001), Nelson et al. (2002), and Dowling (2009), participants completed a field survey that required them to select one recently completed audit for which a fraud brainstorming session was held. They then responded to questions related to the consideration of fraud based upon their experiences on that engagement. Participants provided measures of fraud risk factors, their overall fraud risk assessment for the engagement, measures of audit responses to the fraud risk assessment, and objective and subjective data regarding the quality of the audit's fraud brainstorming session(s).

Fraud Risk Factors

Table 1 provides descriptive data related to fraud risk factors. Fraud risk factors elicited from participants were measured according to SAS No. 99 and prior research (e.g., Wilks and Zimbelman 2004) that characterize fraud risk factors as incentives, opportunities, or rationalizations. We measure incentives with seven questions related to market and debt incentives and opportunity with five measures of corporate governance quality, internal control strength, and accrual levels. We were unable to find research that suggests rationalization can be measured with multiple items, so we use one measure consistent with the wording in SAS No. 99. To facilitate comparisons across participants, we measure all but one of these items (i.e., public versus private company status) on a scale where 1 = extremely low and 10 = extremely high.

Factor analyses indicate that the items associated with market-related incentives and debt-related incentives load on separate factors. Because participants were allowed to provide data from

[3] Because not all individuals responded to our survey, we examine the potential for nonresponse bias. Oppenheim (1992) recommends comparing data from late respondents to early respondents as a way of assessing this bias. Accordingly, we compare the responses from the first quartile of respondents to those of the last quartile of respondents. There are no statistically significant differences between early and late responses on any of our hypothesized variables.

TABLE 1
Descriptive Statistics: Fraud Risk Factors

Item[a]	Response [n = 179] Mean	Std. Dev.
Market Incentive		
Percentage of publicly traded clients	67.03	(47.00)
Incentives or pressures for client to commit fraud	4.48	(2.14)
Pressure for client to meet earnings estimates	4.89	(2.57)
Extent of equity-based management compensation	4.28	(2.52)
Extent of budget-based management compensation	4.68	(2.40)
Debt Incentive		
Pressure for client to meet debt covenants	3.72	(2.68)
Amount of debt	4.93	(2.75)
Opportunity		
Opportunity for fraud to be perpetrated at client	4.90	(1.92)
Strength of corporate governance	6.85	(2.26)
Likelihood of management override of controls	4.32	(2.12)
Overall internal control effectiveness	6.74	(2.05)
Difference between net income and cash flow from operations	4.58	(2.34)
Rationalization		
Extent to which management is enabled to rationalize fraud	3.57	(2.07)

[a] With the exception of whether the client chosen was publicly traded (coded 1) or private (coded 0), all items are measured on a scale where 1 = extremely low and 10 = extremely high. Also, due to the positive wording of two opportunity items (strength of corporate governance and overall internal control effectiveness), responses for these two items are reverse-coded for data analysis.

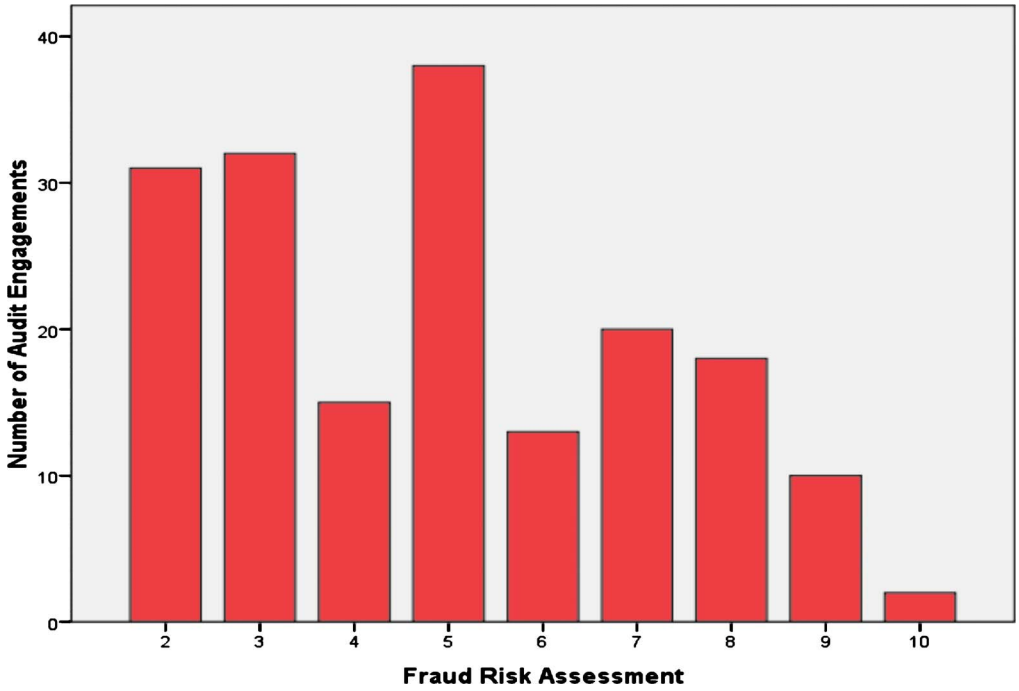
audits of both publicly traded and privately held companies, we decompose the incentive construct into "market incentive" and "debt incentive." We measure the four fraud risk factors as the mean of the participant's responses to the items classified under each factor.[4]

Fraud Risk Assessment and Fraud Risk Responses

Similar to Carpenter (2007), participants provided the overall fraud risk assessment for the engagement on a scale where 1 = extremely low and 10 = extremely high. The non-tabulated mean fraud risk assessment (standard deviation) for the sample was 4.92 (2.23). Figure 2 provides a frequency distribution of fraud risk assessments and illustrates substantial variation among the risk assessments, which ranged from 2 to 10 for our sample. After providing risk assessments, participants were asked to provide data on how the engagement team responded to the assessment by modifying the nature, staffing, timing, and extent of subsequent audit procedures. Table 2

[4] With only three exceptions, factor analyses of items combined to create fraud risk factors and fraud risk responses provide factor loadings in excess of 0.50 and Cronbach's alpha levels exceeding the generally accepted threshold of 0.70 (Nunnally 1978). For all combined measures, the percents of variances explained exceed 60.0 percent. With respect to the exceptions: (1) the factor loading for a measure of total accruals (0.283) was unacceptable and therefore excluded from subsequent analyses, (2) the remaining four measures of opportunity provide a moderate reliability level (Cronbach's alpha = 0.62) that is acceptable for testing relationships (e.g., Bamber et al. 1989), and (3) the Cronbach's alpha level for staffing (0.423) was low and we therefore provide results for staffing with the two staffing items combined and separated.

FIGURE 2
Frequency Distribution of Fraud Risk Assessments



This figure illustrates fraud risk assessments across the sample of actual audit engagements. Fraud risk assessment is measured on a scale ranging from 1 ("extremely low") to 10 ("extremely high").

provides descriptive statistics on the items used to measure fraud risk responses. The "nature" of responses was measured with two items: (1) the degree to which the number of testing procedures was increased, and (2) the degree to which the types of procedures were changed. "Staffing" was determined by both the degree to which more experienced staff or forensic specialists were used on the engagement. The "timing" of audit responses was measured by the degree to which the timing of audit procedures was changed on the engagement. The "extent" of auditors' responses was measured with three items: the extent to which (1) budgeted audit hours, (2) sample sizes, and (3) audit documentation review were increased. All eight items were measured on a scale where 1 = none and 10 = extensively. We measure the four fraud risk responses as the mean of the participant's answers to the underlying items.

Quality of the Fraud Brainstorming Session

We develop a 21-item measure to determine the quality of brainstorming sessions. Table 3 provides descriptive statistics for the 21 items used to measure quality based on three elements: attendance and communication, brainstorming structure and timing, and engagement team effort. Responses for each of the items were coded either 1 (high quality) or 0 (low quality), based upon the participant's response. Similar to DeFond et al.'s (2005) multiple measure of corporate gov-

TABLE 2
Descriptive Statistics: Fraud Risk Responses

Item[a]	Response [n = 179] Mean	Std. Dev.
Nature		
Increase the number of audit procedures	5.26	(2.43)
Change the types of audit procedures	5.56	(2.41)
Staffing		
Use more experienced auditors to perform audit procedures	4.36	(2.78)
Increase the use/consultation of forensic specialists	2.79	(2.73)
Timing		
Change the timing of audit procedures	3.83	(2.70)
Extent		
Increase budgeted hours	3.35	(2.43)
Increase sample sizes	3.73	(2.56)
Increase the extent of audit documentation review	4.71	(2.80)

[a] All items are measured on a scale where 1 = none and 10 = extensively.

ernance quality, we weigh each item equally to avoid adding additional subjectivity to the measure. Thus, we rate the total quality of a brainstorming session on a scale from 0–21. The nontabulated mean quality score for our sample was 10.56 (SD = 3.85), with a minimum score of 3 and a maximum score of 18. Figure 3 provides a frequency distribution of the brainstorming session quality scores, indicating substantial variation in brainstorming quality.

Attendance and Communication

Our measure includes 11 items related to which professionals led/attended the brainstorming sessions and the extent of their contribution. For example, proper tone at the top and a highly qualified person leading the session are important to the success of the brainstorming process (AICPA 2003; Landis et al. 2008). Further, whole engagement team participation allows more experienced auditors to share their fraud insights/experiences with less experienced team members (e.g., AICPA 2003). Therefore, if the brainstorming session is led by either the engagement partner or a fraud specialist, the engagement receives a coding of 1 (i.e., high quality) for this item. We measure the respondent's perceived level of contribution to the brainstorming session from staff, seniors, managers, partners, and various specialists, as well the level of the engagement management's openness to all ideas submitted during the brainstorming session. These items are measured on a scale from 1 (extremely low) to 10 (extremely high) and responses related to these items are coded high if they exceed the sample mean for the item (means provided in Table 3).

Brainstorming Structure and Timing

The survey contains four measures of brainstorming structure and timing. For example, practitioner guidance (e.g., AICPA 2003; Beasley and Jenkins 2003) suggests that the use of agendas should increase the effectiveness and efficiency of brainstorming sessions. Conversely, Asare and Wright (2004) find that the use of fraud checklists can impair auditor performance. Therefore, responses are coded as high quality if an agenda (checklist) is (is not) used.

TABLE 3

Descriptive Statistics: Quality of Fraud Brainstorming Session (21 Items)

Item	Response [n = 179] Mean	Std. Dev.
Attendance and Communication		
1. % of engagements where the session was led by the partner or forensic specialist	59.78	(49.22)
2. % of engagements where all levels of engagement team attended sessions	72.63	(44.50)
3. % of engagements where a fraud specialist attended the primary session	31.29	(46.34)
4. % of engagements where an IT audit specialist attended the primary session	69.27	(46.34)
5. % of engagements where a tax professional attended the primary session	62.56	(48.59)
6. Level of staff contribution to the session [n = 131][a,b]	4.32	(2.13)
7. Level of senior contribution to the session [n = 170][a,b]	6.03	(1.91)
8. Level of manager contribution to the session[a]	7.93	(1.48)
9. Level of partner contribution to the session[a]	8.41	(1.48)
10. Level of fraud specialist contribution to the session [n = 56][a,b]	7.01	(2.31)
11. Level of openness to ideas submitted during the session[a]	8.54	(1.42)
Brainstorming Structure and Timing		
12. % of engagements where an agenda was used during the session	83.80	(36.60)
13. % of engagements where a checklist was not used during the session	28.49	(45.34)
14. % of engagements where the primary session was held pre-planning or early in planning	64.81	(47.95)
15. % of engagements where a session was held at the end of the prior year's audit	16.20	(36.03)
Engagement Team Effort		
16. Time spent by the engagement team preparing for the session (in hours)	9.10	(15.51)
17. % of engagements that had more than one session	49.16	(50.13)
18. Total length of time for the sessions (in minutes)	93.68	(99.16)
19. % of engagements where team members were asked to identify risks prior to the session	67.60	(47.00)
20. Extent of discussion about how management might perpetrate fraud[a]	7.25	(1.69)
21. Extent of discussion about audit responses to fraud risk[a]	7.38	(1.60)

[a] Items are measured on a scale where 1 = extremely low and 10 = extremely high.

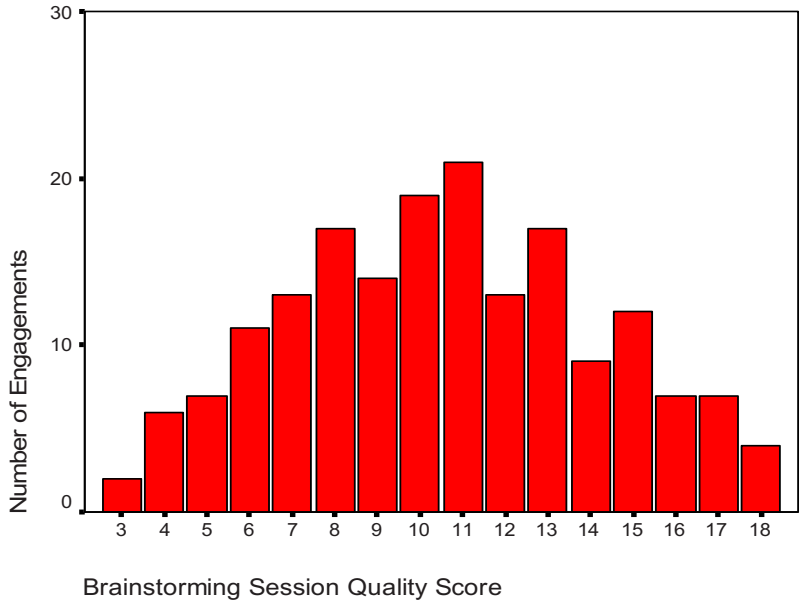
[b] Mean represents the average response in sessions where the professional(s) (e.g., staff, fraud specialist) was present (versus not present). Manager(s) and partner(s) were present for all observations.

Engagement Team Effort

We measure engagement team effort with six items. To illustrate, we deem responses as high if the time spent by the engagement team preparing for the brainstorming session exceeds the mean for our sample, if the engagement has more than one brainstorming session, or if the total length of time for the brainstorming session(s) exceeds the mean for our sample.[5]

[5] Factor analyses performed on the 21 categorical response items provide significant factor loadings ($p < 0.05$) for 20 of the 21 items on their respective constructs. Further, each of these 20 response items provide significant factor loadings ($p < 0.05$) on the overall construct of brainstorming quality. Both the Comparative Fit Index of 0.93 and the Tucker-Lewis Index of 0.93 are above 0.90, which is indicative of a good fit for our overall construct of brainstorming quality (Kline 2004). Only one item, percentage of engagements where all levels of the engagement team attended sessions, did

FIGURE 3
Frequency Distribution of Brainstorming Session Quality Scores



This figure illustrates brainstorming session quality scores across the sample of actual audit engagements. Brainstorming quality scores are formed using a 21-item measure with each item being equally weighted to determine a session's quality. Consequently, session quality scores can range from 0 to 21 points.

Control Variables

Table 4 provides descriptive statistics for our control variables. Prior research has found that client size, industry, and complexity impact auditor risk assessments and testing decisions (e.g., O'Keefe et al. 1994). We control for size by including a measure of the client's revenues in our analyses. We asked respondents to provide the client's industry and, similar to Johnstone and Bedard (2003), create dichotomous industry control variables to control for industry effects. As suggested by the audit fee and resource allocation literatures (e.g., Hackenbrack and Knechel 1997), we measure client complexity as the number of financial statement auditors assigned to the engagement divided by our size measure. Because we investigate fraud-related judgments, we control for whether fraud was detected on the engagement. We also control for the effects of other auditor or engagement team-related variables that could influence audit judgments (e.g., Bonner 1990). Specifically, with dichotomous firm control variables, we simultaneously control for the specific audit firm which performed the audit and the inclusion of one non-Big 4 firm in our sample. Also, our model includes the participant's general audit experience, position in the firm,

not load satisfactorily ($p = 0.70$). Omitting this item from the factor analyses and our session-quality measure in our regression analyses does not qualitatively change our results. We include this item in our analyses because the item was specifically identified as a deficiency by the PCAOB (PCAOB 2007).

TABLE 4

Descriptive Statistics: Control Variables

	Response [n = 179]	
Item[a]	**Mean**	**Std. Dev.**
Client Size	3.06	(1.42)
Manufacturing industry	0.20	(0.40)
Retail industry	0.11	(0.31)
Government/not-for-profit industry	0.19	(0.39)
Energy industry	0.08	(0.28)
High tech/communications industry	0.14	(0.35)
Healthcare/pharmaceutical industry	0.07	(0.26)
Financial services industry	0.07	(0.25)
Miscellaneous industry	0.14	(0.35)
Client complexity	0.93	(0.65)
Fraud detect	0.24	(0.42)
Audit firm A	0.14	(0.35)
Audit firm B	0.19	(0.39)
Audit firm C	0.32	(0.47)
Audit firm D	0.16	(0.37)
Audit firm E	0.19	(0.39)
Audit experience	1.70	(0.86)
Position	2.29	(1.23)
Client experience	3.35	(2.05)
Industry experience	8.69	(6.73)
Engagement team expertise	7.50	(1.81)
Fraud training	1.42	(0.66)
Fraud experience	1.92	(0.71)
Restart	0.08	(0.28)
Time taken	22.29	(15.51)

[a] Items were measured as follows:

Client size is based on client revenues and coded as follows: 1 = < $100 million, 2 = $100 million −$500 million, 3 = > $500 million −$1 billion, 4 = > $1 billion −$5 billion, 5 = > $ 5 billion. Industry variables are dichotomous (1,0) industry variables for the following industries: manufacturing, retail, government/not-for-profit, energy, high tech/communications, healthcare/pharmaceuticals, financial services, and miscellaneous. Client complexity is measured as number of financial statement auditors assigned to the engagement (coded as follows: 1 = 0–5 auditors, 2 = 6–10 auditors, 3 = 11–15 auditors, 4 = 16–20 auditors, 5 = > 20 auditors)/client size. Fraud detect is coded 1 if the participant reported that fraud was detected at the client, 0 otherwise. Audit firm variables are dichotomous (1,0) firm variables for the four international and one national (i.e., non-Big 4) firm that participated in our study. Audit experience is coded as follows: 1 = 5–10 years, 2 = 11–15 years, 3 = > 15 years. Position is coded as follows: 1 = manager, 2 = senior manager, 3 = director, 4 = partner. Client experience is measured as the number of years the respondent had served the client. Industry experience is coded as follows: 1 = 0–5 years, 2 = 6–10 years, 3 = 11–15 years, 4 = 16–20 years, 5 = > 20 years. Engagement team expertise is measured on a scale where 1 = extremely low and 10 = extremely high. Fraud training is measured by the number of hours of fraud training the respondent had received in the last year and coded as follows: 1 = 0–8 hours, 2 = 9–16 hours, 3 = > 16 hours. Fraud experience is determined by the number of engagements the respondent served on which frauds were identified and coded as follows: 1 = 0, 2 = 1–2, 3 = >2. Restart is coded 1 if the participant stopped and restarted the survey, 0 otherwise. Time taken is the time taken by the 164 respondents who did not stop and restart the survey.

experience with the client and the client's industry, extent of fraud training, experience level with actual financial statement fraud, their perception of the entire engagement team's expertise level, whether the respondent stopped and started the survey, and time taken to complete the survey.

Regression Models

To test H1 and H2, we estimate the following models via ordinal logistical regression and examine the italicized interactions:

H1:

$$FR = \beta_0 + \beta_1 MI + \beta_2 DI + \beta_3 OPP + \beta_4 RAT + \beta_5 SESSION\ QUALITY + \beta_6 MI \times SESSION\ QUALITY + \beta_7 DI \times SESSION\ QUALITY + \beta_8 OPP \times SESSION\ QUALITY + \beta_9 RAT \times SESSION\ QUALITY + \beta_{10-32}\text{CONTROL VARIABLES} + \varepsilon.$$

H2:

$$NATURE = \beta_0 + \beta_1 FR + \beta_2 SESSION\ QUALITY + \beta_3 FR \times SESSION\ QUALITY + \beta_{4-26}\text{CONTROL VARIABLES} + \varepsilon.$$

$$STAFFING = \beta_0 + \beta_1 FR + \beta_2 SESSION\ QUALITY + \beta_3 FR \times SESSION\ QUALITY + \beta_{4-26}\text{CONTROL VARIABLES} + \varepsilon.$$

$$TIMING = \beta_0 + \beta_1 FR + \beta_2 SESSION\ QUALITY + \beta_3 FR \times SESSION\ QUALITY + \beta_{4-26}\text{CONTROL VARIABLES} + \varepsilon.$$

$$EXTENT = \beta_0 + \beta_1 FR + \beta_2 SESSION\ QUALITY + \beta_3 FR \times SESSION\ QUALITY + \beta_{4-26}\text{CONTROL VARIABLES} + \varepsilon.$$

where:

FR = Fraud Risk Assessment, measured on a scale where 1 = extremely low and 10 = extremely high;

MI, *DI*, *OPP*, *RAT* = Market Incentive, Debt Incentive, Opportunity, and Rationalization; see the "Fraud Risk Factors" section for information on measurement;

SESSION QUALITY = see the "Quality of the Fraud Brainstorming Session" section for information on measurement; this variable is coded 1 if session quality for the engagement exceeds the mean for the sample (10.56), 0 otherwise; the use of an indicator variable is described in the "Correlation Matrix and Multicollinearity" section below;

NATURE, *STAFFING*, *TIMING*, *EXTENT* = see the "Fraud Risk Assessment and Fraud Risk Responses" section for information on measurement;

CONTROL VARIABLES (*CLIENT SIZE*, *INDUSTRY VARIABLES*, *CLIENT COMPLEXITY*, *FRAUD DETECT*, *AUDIT FIRM VARIABLES*, *AUDIT EXPERIENCE*, *POSITION*, *CLIENT EXPERIENCE*, *INDUSTRY EXPERIENCE*, *ENGAGEMENT TEAM EXPERTISE*, *FRAUD TRAINING*, *FRAUD EXPERIENCE*, *RESTART*, *TIME TAKEN*) = see the "Control Variables" section and Table 4 for information on measurement.

IV. RESULTS

Descriptive Statistics of Brainstorming Practices

Given our sample of 179 recently completed audit engagements across five public accounting firms, Table 3 provides descriptive evidence regarding how selected fraud brainstorming sessions are being performed in practice. First, our participants report that a partner or forensic specialist led the session only 60 percent of the time, pointing toward the important role audit managers and seniors probably play in determining the quality of the session. Second, consistent with recent PCAOB inspection team reports (PCAOB 2007), we find evidence that the attendance of specialists (fraud, tax, IT) in sessions varies. Third, our results are consistent with the hierarchical structure of engagement teams in that the level of contribution to the session increases with seniority. Fourth, consistent with PCAOB inspection team findings (PCAOB 2007), for 35 percent of engagements, sessions were held *after* the early stages of planning. Last, engagement team preparation for the sessions tended to be substantial, sessions lasted an average of 1.5 hours, and 49 percent of engagements conducted more than one session.

Correlation Matrix and Multicollinearity

Table 5 presents a correlation matrix. To present a parsimonious correlation matrix, brainstorming quality items and control variables are excluded from Table 5 if they are not significantly correlated ($p < 0.05$) with at least four of the hypothesized variables.

We test our hypotheses via interactions within five ordinal logistical regression models. These interactions are the cross product term of two independent variables (e.g., fraud risk $\times$ session quality). In such regressions, the interactive term is likely to covary with the main effects constituting the interaction term. These regressions often exhibit multicollinearity problems (Echambadi and Hess 2007). Likewise, when we interact our 21-item measure for session quality with risk factors (H1) and risk assessments (H2), diagnostics suggest multicollinearity exists between the interaction terms and the variables that constitute the interaction (e.g., variance inflation factors [VIFs] in excess of 10, condition indices in excess of 30 with more than 50 percent of the variance of coefficient estimates associated with these indices). Multicollinearity, similar to a small sample size, biases against finding statistically significant main effects and interactions in multivariate regression (e.g., Belsley et al. 1980; Hair et al. 1998; Wooldridge 2009).

Prior studies have mitigated multicollinearity effects by transforming their continuous or scaled moderating variable into an indicator variable (e.g., Core et al. 2002; Ali et al. 2008; Cardinaels 2008). Thus, we transform our 21-item measure of brainstorming quality into an indicator variable, coding each engagement's total session quality as high or 1 (low or 0) if the total score for the engagement exceeds (is less than) the mean for our sample (versus using the 21-item measure).[6] We use this indicator variable to examine the main and interactive effects of session quality in all of our tabulated and non-tabulated regression analyses. Tests of multicollinearity for our H1 regression analyses (with the session quality indicator variable) indicate that all but one of the independent variables for H1 provide satisfactory multicollinearity diagnostics (e.g., Belsley et al. 1980; Neter et al. 1996; Kennedy 2008). The sole exception is the VIF for the interaction between opportunity and session quality, which is 11.71. Using condition indices and

[6] Alternatively, we could have split the sample at the median session quality score (11.00) and excluded firms at the median from our analyses. However, we would then have eliminated 21 observations with a session quality of 11.00 from our analyses. Given our relatively small sample size (179), our large number of control variables (23), that the mean and median for our sample are approximately the same (10.56 and 11.00, respectively), and partitioning the sample at the mean can be performed without eliminating observations (i.e., no observations had a session quality score = 10.56), we use the mean session quality score to create the indicator variable for session quality.

TABLE 5

Correlation Matrix

		MI[a]	***DI***	***OPP***	***RAT***	***FR***	***NATURE***	***STAFFING***	***TIMING***	***EXTENT***
DI[a]	Correlation[b]	***0.22***								
OPP[a]	Correlation	0.03	***0.29***							
RAT[a]	Correlation	***0.22***	0.13	***0.42***						
FR[a]	Correlation	***0.24***	***0.22***	***0.32***	***0.31***					
NATURE[a]	Correlation	***0.23***	0.09	0.07	***0.17***	***0.39***				
STAFFING[a]	Correlation	***0.19***	0.13	0.06	***0.16***	***0.36***	***0.61***			
TIMING[a]	Correlation	***0.22***	0.10	0.13	***0.20***	***0.31***	***0.45***	***0.43***		
EXTENT[a]	Correlation	***0.16***	0.09	0.13	***0.20***	***0.46***	***0.69***	***0.64***	***0.50***	
PTR	Correlation	0.08	0.13	0.02	0.10	0.14	0.15	***0.25***	***0.19***	***0.18***
IT	Correlation	***0.18***	−0.07	***−0.29***	−0.06	0.00	0.09	***0.18***	***0.17***	0.10
EARLY	Correlation	0.00	0.08	−0.01	−0.05	***0.15***	***0.18***	***0.15***	***0.19***	***0.17***
FD	Correlation	***0.20***	−0.02	***−0.15***	0.10	***0.16***	***0.24***	***0.29***	0.11	***0.19***
RD	Correlation	***0.18***	−0.02	***−0.17***	0.08	0.13	***0.31***	***0.26***	***0.17***	***0.28***
MC	Correlation	0.13	−0.06	−0.12	−0.04	***0.17***	***0.25***	***0.24***	***0.15***	***0.24***
PC	Correlation	0.01	−0.05	***−0.16***	−0.03	0.14	***0.19***	***0.20***	***0.15***	0.14
SQ[a]	Correlation	***0.19***	0.00	***−0.16***	0.05	0.13	***0.30***	***0.40***	***0.28***	***0.30***
EXP	Correlation	***0.16***	−0.05	***−0.29***	−0.04	0.06	***0.22***	***0.21***	0.12	0.06
SIZE	Correlation	***0.27***	−0.01	***−0.18***	−0.12	−0.08	0.12	***0.26***	***0.14***	0.06
		PTR	***IT***	***EARLY***	***FD***	***RD***	***MC***	***PC***	***SQ***	***EXP***
IT	Correlation	***0.29***								
EARLY	Correlation	***0.23***	0.07							
FD	Correlation	0.07	0.06	0.03						
RD	Correlation	0.11	0.07	0.01	***0.74***					
MC	Correlation	−0.13	−0.03	0.03	***0.34***	***0.43***				
PC	Correlation	0.03	−0.04	0.06	***0.35***	***0.32***	***0.52***			
SQ[a]	Correlation	***0.36***	***0.37***	***0.29***	***0.51***	***0.48***	***0.45***	***0.42***		
EXP	Correlation	−0.02	***0.19***	−0.01	***0.24***	***0.20***	***0.18***	***0.16***	***0.22***	
SIZE	Correlation	0.11	***0.32***	−0.05	***0.17***	***0.17***	0.11	0.09	***0.31***	***0.32***

(continued on next page)

TABLE 5 (continued)

Correlations in bold italics are significantly correlated (p-value < 0.05).

[a] The "Method" section provides information on the measurement of market incentive (*MI*), debt incentive (*DI*), opportunity (*OPP*), rationalization (*RAT*), fraud risk assessment (*FR*), *NATURE*, *STAFFING*, *TIMING*, *EXTENT*, and brainstorming session quality (*SQ*). For illustrative purposes, the scaled (1–21) version of the *SQ* variable is presented in this table.

[b] Pearson correlation statistic.

Variable Definitions:

PTR = whether the session was led by the partner or forensic specialist (coded = 1, 0 otherwise);
IT = whether an IT audit specialist attended the primary session (coded = 1, 0 otherwise);
EARLY = whether the engagement's primary session was held pre-planning or early in planning (coded = 1, 0 otherwise);
FD = extent of discussion about how management might perpetrate fraud, measured on a scale where 1 = extremely low and 10 = extremely high;
RD = extent of discussion about audit responses to fraud risk, measured on a scale where 1 = extremely low and 10 = extremely high;
MC = level of manager contribution to the session, measured on a scale where 1 = extremely low and 10 = extremely high;
PC = level of partner contribution to the session, measured on a scale where 1 = extremely low and 10 = extremely high;
EXP = engagement team expertise, measured on a scale where 1 = extremely low and 10 = extremely high; and
SIZE = client size based on client revenues and coded as follows: 1 = < $100 million, 2 = $100 million–$500 million, 3 = > $500 million–$1 billion, 4 => $1 billion–$5 billion, 5 => $5 billion.

variances to assess multicollinearity (e.g., Belsley et al. 1980; Core et al. 2002), we find that only one of the 33 condition indices for H1 exceeds the standard threshold of 30, and over 50 percent of the variances for the intercept and *session quality* are associated with this index (Belsley et al. 1980; Kennedy 2008). This is likely the result of session quality being included in the model for H1 five times (one main effect and four interactions).[7] Thus, for testing H1, multicollinearity is likely to bias against finding significant main and interactive effects for *session quality,* and caution must be taken in interpreting null findings for this variable. For H2 analyses (with the session quality indicator variable), both VIFs and condition indices/variances indicate no problems with multicollinearity, as all measures are below acceptable thresholds (Neter et al. 1996; Belsley et al. 1980; Kennedy 2008).[8]

Hypotheses Testing

Fraud Risk Factors, Fraud Risk Assessments, and Brainstorming Session Quality

Hypothesis 1 posits that fraud risk factors are more positively related to fraud risk assessments in high-quality brainstorming sessions than in low-quality sessions. Thus, H1 is supported by a regression result that provides positive and significant interactions between fraud risk factors and session quality on fraud risk assessments. Table 6 provides the results of H1 testing. For parsimony, only control variables with p-values < 0.10 are included in Tables 6 and 7. While the main effects of market incentive, opportunity, and rationalization are all positive and significant (p-values < 0.05), their interactions with brainstorming session quality are not. Thus, while we find that three of the four risk factors are positively associated with fraud risk assessments, we do not observe that these relations are stronger under higher quality brainstorming.

However, diagnostics suggest that multicollinearity influences the main and interactive effects of session quality in the regression model for H1 (specifically related to its interaction with opportunity). As suggested by Bradshaw et al. (2004) and Covrig et al. (2007), we complement our analysis of H1 with two separate regressions with session quality as low and high (two subsamples). We partition the sample at the mean brainstorming quality score (10.56), and re-perform the statistical tests described above for each of the two groups (removing the brainstorming quality variable). If session quality positively moderates the effect of fraud risk factors on fraud risk assessments, then the main effects of fraud risk factors on fraud risk assessments should be more significant for the high-quality subsample. We find this to be the case for the fraud risk factors of debt incentive and opportunity. For the high-quality group, tests of H1 (not tabulated) indicate at least marginally significant and positive relations between debt incentive and opportunity on fraud risk assessments (p-values $= 0.06$ and 0.04, respectively). For the low-quality group, the relations between debt incentive and opportunity on fraud risk assessments are not significant. Thus, there is limited evidence supporting H1.[9]

---

[7] Given the significant correlations between our fraud risk factors (see Table 5) and that auditors *simultaneously* incorporate fraud risk factors into one fraud risk assessment, not including all four fraud risk factors in the same multivariate regression could yield spurious results.

[8] To ensure that our interaction results are stable across various specifications, we analyze the following models: (1) with both main and interactive effects for continuous hypothesized variables, (2) with both main and interactive effects for continuous hypothesized variables (mean-centered), (3) with only interactive effects for continuous hypothesized variables (no mean-centering), and (4) with both main and interactive effects for continuous hypothesized variables (no mean-centering, but session quality as an indicator variable (the tabulated model)). Our primary inferences are similar across all four models: the interactions for H1 are largely not supported and the interactions for H2 are largely supported. It should be noted that Echambadi and Hess (2007) conclude that mean-centering does not alleviate collinearity problems in multivariate regressions.

[9] Notwithstanding our expectation, we find only limited evidence that the quality of brainstorming sessions moderates the relations between risk factors and assessments. Anecdotal evidence from conversations with partners at the Big 4 firms suggests one possible explanation for our results. These partners report that engagement team management come to

TABLE 6

H1 Testing: Regression Results for Fraud Risk Assessment[a]

Independent Variable[b]	Estimated Coefficient	Wald-Statistic	p-value
MI	0.271	2.96	0.043
DI	0.095	1.07	0.150
OPP	0.317	3.13	0.039
RAT	0.249	4.35	0.019
SESSION QUALITY	0.409	0.11	0.739
MI × *SESSION QUALITY*	0.001	0.00	0.499
DI × *SESSION QUALITY*	0.018	0.02	0.443
OPP × *SESSION QUALITY*	0.086	0.14	0.356
RAT × *SESSION QUALITY*	−0.132	0.71	0.398
CLIENT SIZE	−0.252	3.10	0.078
ENGAGEMENT TEAM EXPERTISE	0.200	3.75	0.053
FRAUD TRAINING	0.649	7.23	0.007
RESTART	−1.580	3.51	0.061
Model Chi-square statistic = 79.60 (p-value = < .001)			
Pseudo R^2 = . 362			

[a] Fraud risk assessment (*FR*) is measured on a scale where 1 = extremely low and 10 = extremely high.

[b] The "Method" section provides information on the measurement of market incentive, debt incentive, opportunity, rationalization. *SESSION QUALITY* is coded 1 if the total session quality of the engagement (described in the "Method" section) exceeded the sample mean (10.56), 0 otherwise. *CLIENT SIZE* is based on client revenues and coded as follows: 1 =< $100 million, 2 = $100 million–$500 million, 3 => $500 million–$1 billion, 4 => $1 billion–$5 billion, 5 => $5 billion. *ENGAGEMENT TEAM EXPERTISE* is measured on a scale where 1 = extremely low and 10 = extremely high. *FRAUD TRAINING* is measured by the number of hours of fraud training the respondent had received in the last year and coded as follows: 1 = 0–8 hours, 2 = 9–16 hours, 3 = > 16 hours. *RESTART* is coded 1 if the participant stopped and restarted the survey, 0 otherwise.

Fraud Risk Assessments, Fraud Risk Responses, and Brainstorming Session Quality

Hypothesis 2 posits that fraud risk assessments are more positively related to fraud risk responses in high-quality brainstorming sessions than in low-quality sessions. Table 7 presents results related to H2. For nature, staffing, timing, and extent, the interaction terms for fraud risk assessment and session quality are all positive and at least marginally significant (p-values < 0.10, see Panels A–D), thus supporting H2. When brainstorming session quality is higher, audit teams are more likely to add or change testing procedures, employ more senior staffing/fraud specialists, change the timing of tests, and increase the extent of testing in response to higher fraud risk

brainstorming sessions with a lot of information about the likelihood of fraud based on prior years' work and client continuation discussions. Specifically, partners indicate that fraud risk assessments are often based on the presence or absence of specific risk factors such that assessments are effectively predetermined by the partners prior to the brainstorming sessions. In such instances, brainstorming serves primarily to inform team members about the known risk factors, to discuss other possible factors that might not have been pre-identified and, more importantly, to give them an opportunity to discuss how best to respond to fraud risk. How previous client experiences, client continuation discussions, and other factors affect the brainstorming process and fraud-related judgments is a question for future research.

TABLE 7

H2 Testing: Regression Results for Fraud Risk Responses[a]

Panel A: Nature

Independent Variable[b]	Estimated Coefficient	Wald-Statistic	p-value
FR	0.229	5.22	0.022
SESSION QUALITY	−0.148	0.05	0.832
FR × SESSION QUALITY	0.196	2.26	0.067
CLIENT SIZE	0.273	3.92	0.048
HEALTHCARE/PHARMACEUTICAL INDUSTRY	−1.903	7.92	0.005
CLIENT COMPLEXITY	0.643	5.68	0.017
FRAUD DETECTION	0.667	3.98	0.046
AUDIT FIRM A	−1.294	6.14	0.013
AUDIT FIRM C	−0.888	3.88	0.049
AUDIT FIRM E	−1.294	6.14	0.013
ENGAGEMENT TEAM EXPERTISE	0.162	3.38	0.066
INDUSTRY EXPERIENCE	0.055	3.42	0.064
Model Chi-square statistic = 87.35 (p-value < .001)			
Pseudo R^2 = .390			

Panel B: Staffing

Independent Variable	Estimated Coefficient	Wald-Statistic	p-value
FR	0.154	2.30	0.130
SESSION QUALITY	−0.130	0.03	0.854
FR × SESSION QUALITY	0.260	3.84	0.025
CLIENT SIZE	0.504	12.52	<0.001
CLIENT COMPLEXITY	0.599	5.00	0.025
FRAUD DETECTION	0.792	5.44	0.020
Model Chi-square statistic = 94.62 (p-value < .001)			
Pseudo R^2 = . 414			

Panel C: Timing

Independent Variable	Estimated Coefficient	Wald-Statistic	p-value
FR	0.101	0.95	0.330
SESSION QUALITY	−0.939	1.71	0.191
FR × SESSION QUALITY	0.328	5.96	0.008
MANUFACTURING INDUSTRY	1.097	4.39	0.036
AUDIT FIRM A	−1.29	5.13	0.023
AUDIT FIRM B	−1.90	13.37	<0.001
CLIENT EXPERIENCE	−0.137	3.19	0.074
FRAUD EXPERIENCE	−0.426	3.23	0.072
Model Chi-square statistic = 78.31 (p-value < .001)			
Pseudo R^2 = . 358			

(continued on next page)

Panel D: Extent

Independent Variable	Estimated Coefficient	Wald -statistic	p-value
FR	0.259	6.64	0.010
SESSION QUALITY	0.095	0.19	0.891
FR × SESSION QUALITY	0.199	2.32	0.064
CLIENT SIZE	0.234	2.92	0.088
RETAIL INDUSTRY	−1.163	3.52	0.061
FRAUD DETECTION	0.920	7.50	0.006
AUDIT FIRM B	−1.094	4.91	0.027
AUDIT FIRM C	−1.072	5.66	0.017
AUDIT FIRM E	−1.225	5.58	0.018
AUDIT EXPERIENCE	−0.763	4.59	0.032
CLIENT EXPERIENCE	−0.130	3.10	0.078
Model Chi-square statistic = 85.52 (p-value < .001)			
Pseudo R^2 = . 383			

[a] The "Method" section provides information on the measurement of the fraud risk responses of *NATURE*, *STAFFING*, *TIMING*, and *EXTENT*.

[b] *FR* is fraud risk assessment, measured on a scale where 1 = extremely low and 10 = extremely high. *SESSION QUALITY* is coded 1 if the total session quality of the engagement (described in the "Method" section) exceeds the sample mean (10.52), 0 otherwise. *CLIENT SIZE* is based on client revenues and coded as follows: 1 =< $100 million, 2 = $100 million–$500 million, 3 => $500 million–$1 billion, 4 => $1 billion–$5 billion, 5 => $5 billion. *HEALTHCARE/PHARMACEUTICAL INDUSTRY* is a dichotomous industry variable coded 1 if healthcare/pharmaceutical industry client, 0 otherwise. *CLIENT COMPLEXITY* is measured as the number of financial statement auditors assigned to the engagement (coded as follows: 1 = 0–5 auditors, 2 = 6–10 auditors, 3 = 11–15 auditors, 4 = 16–20 auditors, 5 = > 20 auditors)/*CLIENT SIZE*. *FRAUD DETECTION* is coded 1 if the participant reported that fraud was detected at the client, 0 otherwise. *AUDIT FIRM* variables are dichotomous (1,0) firm variables for the four international and one national (i.e., non-Big 4) firm that participated in our study. *ENGAGEMENT TEAM EXPERTISE* is measured on a scale where 1 = extremely low and 10 = extremely high. *INDUSTRY EXPERIENCE* is coded as follows: 1 = 0–5 years, 2 = 6–10 years, 3 = 11–15 years, 4 = 16–20 years, 5 = > 20 years. *MANUFACTURING INDUSTRY* is a dichotomous industry variable coded 1 if manufacturing industry client, 0 otherwise. *CLIENT EXPERIENCE* is measured as the number of years the respondent had served the client. *FRAUD EXPERIENCE* is determined by the number of engagements the respondent served on which frauds were identified and coded as follows: 1 = 0, 2 = 1–2, 3 = > 2. *RETAIL INDUSTRY* is a dichotomous industry variable coded 1 if retail industry client, 0 otherwise. *AUDIT EXPERIENCE* is coded as follows: 1 = 5–10 years, 2 = 11–15 years, 3 = > 15 years.

assessments. Consistent with H2, a higher quality brainstorming session appears to substantially improve the risk-based strategy prescribed by SAS No. 99 (i.e., a positive relation between risk assessments and responses).[10]

To further illustrate that relations between fraud risk assessments and responses are stronger when session quality is high, we examine the response of *nature* in cases where risk assessment is "high" (i.e., 7–10 on our ten-point scale or the top quartile of risk assessments). The mean responses under high- and low-quality sessions are 8.19 and 5.37, respectively ($p < 0.01$), suggesting that under-auditing could be occurring when brainstorming quality is low.[11]

[10] As noted in footnote 4, the reliability statistic for staffing was below a satisfactory level. Therefore, we examine the interaction posited by H2 in relation to our two measures of staffing: increase use/consultation of fraud specialists and use of more experienced auditors. Non-tabulated results indicate that the interaction term for fraud risk assessment and brainstorming quality is positive and marginally significant (p-values < 0.10) for both measures.

[11] In order to provide results comparable to prior experimental research examining frauds at public companies (e.g., Carpenter 2007), we examine our hypotheses with only the publicly traded companies in our sample (sample size = 120). Non-tabulated regression results are qualitatively similar to those tabulated in this study.

Supplemental Discussion and Analyses

Specific Brainstorming Quality Items Driving Effects

While our 21-item measure of brainstorming quality is motivated by the literature and appears to improve the responsiveness of fraud judgments, it might not be practical for audit teams to adjust all 21 items at once to improve their brainstorming sessions. We therefore examine if a smaller number of items could be identified that would produce the same moderating effects as our 21-item measure. As the correlation matrix (Table 5) indicates, it appears that seven brainstorming items are significantly correlated with a number of the hypothesized variables in this study. These items are: (1) whether the session was led by the partner or forensic specialist, (2) whether an IT audit specialist attended the primary session, (3) whether the engagement's primary session was held pre-planning or early in planning, (4) the extent of discussion about how management might perpetrate fraud, (5) the extent of discussion about audit responses to fraud risk, (6) the level of manager contribution to the session, and (7) the level of partner contribution to the session. These are all controllable inputs that can be easily fostered by engagement management.

In testing H1 and H2, we replace the brainstorming quality indicator variable based upon 21 items with one based upon these seven items. Non-tabulated regression results reveal that this seven-item measure positively moderates the effect of rationalization on fraud risk assessments ($p = 0.03$). Recall that in Table 6, using a session-quality measure based on 21 items, all interactions between fraud risk factors and session quality are not significant. Similarly, the indicator variable for session quality based upon seven items positively moderates the relations between fraud risk assessments and the response factors of nature and staffing (p-values < 0.05). It is especially promising that concentrating on the aforementioned seven items could improve the often tenuous link between fraud risk assessments and the nature of testing. While our results using only the seven items do not provide exactly the same results as those tabulated with the 21-item measure, it appears that ensuring these seven items occur is a practical first step, or set of best practices, to improve the quality of brainstorming.[12]

Control Variables

With respect to the direct effects of control variables on fraud risk assessment in Table 6, we observe a negative relation between client size and fraud risk assessments. Consistent with Beasley et al. (1999), who find that fraud can be more prevalent at smaller companies; auditors tend to assess fraud risk higher for smaller clients. We find that engagement team expertise has a positive association with risk assessments, suggesting that auditors with more expertise are assigned to riskier engagements. We also find a positive effect for fraud training on risk assessment. Finally, participants who stopped and restarted the survey provided lower fraud risk assessments than other participants. Given this result, we re-perform our analyses of H1 and H2 without the participants who stopped and restarted and observe qualitatively similar results.[13]

The mean fraud risk assessment for our sample (4.92 on a ten-point scale) could be considered high, as the base rate for fraud is fairly low. On the other hand, given the relatively high rate of fraud detection in our sample (24 percent), a higher mean fraud risk assessment might be expected. In fact, the mean fraud risk assessment for the subset of audits where fraud was detected ($n = 43$) is a surprisingly low 5.44, and is only marginally higher than the mean for the non-fraud subset (mean $= 4.75$; $p = 0.09$; non-tabulated t-test). Moreover, fraud detection is not signifi-

[12] When we perform the same analyses with the remaining 14 items, there are no significant results for H1 or H2. This suggests that the seven items described above are driving our results.

[13] It is unclear why these participants provided lower fraud risk assessments. A review and analysis of the data for these participants reveal no clear explanation of the variable's significance.

cantly associated with fraud risk assessment in Table 6 (non-tabulated $p > 0.50$). Thus, while auditors in our sample were generally adept at incorporating fraud risk factors into fraud risk assessments, given *ex post* evidence of fraud detection, our results suggest that auditors still could improve the accuracy of their fraud risk assessments.

With respect to fraud responses, Table 7 illustrates that, as client size increases, auditors appear to be more likely to adjust the nature, staffing, and extent of testing. The larger audit fees/budgets for bigger clients typically allow for more testing when needed. Larger clients also tend to be more prestigious and, if needed to respond to fraud risks, engagement leaders can typically acquire additional audit professionals for the engagement. The negative relation between the healthcare/pharmaceutical industry and the nature of testing likely reflects that this industry was the only industry in our sample not to have fraud detected during the audit (possibly due to the industry being heavily regulated).

As one would expect, as client complexity increases, audit teams tend to alter the nature of testing and use more competent staff. Of particular interest is the positive link between the detection of fraud and the nature, staffing, and extent of procedures. These responses appear to be integral to detecting fraud, and it could be important for auditors to adjust their testing in multiple ways to improve fraud detection. It is also interesting to note that fraud detection is not significantly associated with the timing of tests. Consistent with this finding, we observe negative relations between both client experience and fraud experience and the timing of testing. These combined findings suggest that experienced auditors likely question the effectiveness of changing the timing of testing to detect fraud. However, this relation could be industry specific, as we find a positive association between the manufacturing industry and the timing of tests. Given the risks associated with manipulating ending inventory, auditors of manufacturing clients appear to be more apt to change the timing of their inventory examinations.

We find that audit firms are related to a number of our responses. Firms have different audit approaches or structures, and these differences can lead to different audit judgments (Prawitt 1995; Hyatt and Prawitt 2001). As one would expect, higher levels of engagement team expertise and industry experience are positively associated with the difficult task of changing the nature of testing. Last, we observe negative associations between both audit experience and client experience and the extent of the response. Given the positive relationship between fraud detection and the extent of testing (i.e., evidence that changing the extent of procedures can be an effective fraud detection tool), these combined findings suggest a negative effect of longer auditor tenure on audit quality. Such relations could, to some extent, be mitigated by the Sarbanes-Oxley Act of 2002, which mandates a five-year rotation for the lead and concurring audit partners on publicly traded clients (U.S. House of Representatives 2002).

Other Data Issues

As previously noted, fraud was detected on 24 percent of the engagements in our sample, a much higher rate than reported by previous studies (e.g., Loebbecke et al. 1989; Nieschwietz et al. 2000). This high rate of fraud could be due to the fact that fraud cases were more salient to participants when they selected an audit as the basis for responding to our survey. We asked for detailed information on the types and sizes of the frauds that were detected. However, few participants chose to provide this information. One respondent reported a fraud related to management's manipulation of payroll reports, while another reported a fraud related to payments to a fictitious vendor. In comparison to the other industries in our sample, both retail and governmental/not-for-profit entities had unusually high rates of fraud detection (36 percent and 32 percent, respectively). In retail, the misappropriation of retail inventory, whether material or not, could have contributed to its high rate. For governmental and not-for-profit entities, the incentive to fraudulently misstate financial statements is lower than in other industries. Thus, it is possible

that some frauds in our sample were related to misappropriations of assets (versus fraudulent financial reporting) and were likely immaterial from the auditor's perspective.

Given the high rate of fraud in our sample, we consider whether our primary results hinge on recalled audits that detected fraud. We examine this concern in two ways. First, we control for whether fraud was detected in all of our analyses. We find support for H2 even though the fraud detection control variable is significant in several of our analyses (see Table 7, Panels A, B, and D). Second, in non-tabulated analyses, we remove the 43 engagements where fraud was detected and find qualitatively similar results (i.e., H1 (H2) is not (is) supported).[14] Thus, our results do not appear to be an artifact of engagements for which fraud was detected.

We also examine whether our brainstorming data presented in Table 3 are significantly skewed by these engagements. Non-tabulated t-tests for each of the 21 items find that none of the means are significantly different between audits where fraud was or was not detected. Likewise, correlations between the fraud detection control variable and the 21-item measure and indicator versions of brainstorming session quality are not significant (p-values > 0.50). Last, the mean brainstorming quality score for the fraud sample (10.56) is actually equal to the mean for our entire sample (10.56) and is not significantly different than the mean for the non-fraud sample (10.57; $p > 0.50$). Thus, it appears that the descriptive brainstorming data presented in Table 3 is not significantly skewed by our fraud observations and that, for audits where fraud was detected, participants were not biased toward providing responses reflecting high-quality brainstorming sessions. In summary, these additional analyses suggest that our results are not primarily attributable to the audits in our sample where fraud was detected.

From an association perspective, it might be that audits with higher fraud risk assessments prompt auditors to conduct higher quality brainstorming sessions and carefully link fraud judgments as precautionary measures. Consequently, our results might not be reflective of audits with lower fraud risk assessments. In Table 5, we find that session quality is only marginally correlated with fraud risk assessments ($p = 0.09$) and, in a multivariate setting (Table 6), we find that fraud risk assessments and session quality are not significantly related ($p > 0.50$). Still, we re-perform our regressions with only audits where the fraud risk assessment ranged from 2–6 (removing engagements with the highest fraud risk assessments (7–10 or the top quartile of fraud risk assessments)). As noted in Figure 2, no respondents provided a fraud risk assessment of 1. Non-tabulated results related to this subset of less risky audits are qualitatively similar to those tabulated. Thus, our results do not appear to be driven by the riskier audit engagements in our sample.

V. CONCLUSION

SAS No. 99 requires brainstorming sessions on every audit (AICPA 2002), but the PCAOB has raised concerns about the quality of these sessions in practice (PCAOB 2007). This is the first study, to our knowledge, to consider whether and how the *quality* of brainstorming influences team decision-making.

Our study makes several contributions to research and practice. Leveraging the field survey method, we develop and test a measure of brainstorming quality using data collected from 179 actual audit engagements. We find that fraud risk factors are positively correlated with fraud risk assessments, and we find some evidence that brainstorming quality moderates these relations. More importantly, we find that when brainstorming quality is perceived to be higher, risk assessments are more positively related to the nature, staffing, timing, and extent of audit procedures. Achieving a high level of brainstorming quality appears to improve the audit team's consideration

[14] In tests of H1, excluding the 43 engagements where fraud was detected, all independent variables of interest are not significant (p-values > 0.50). In tests of H2, all independent variables of interest are significant (p-values < 0.10).

of fraud by cultivating a broader set of responses to identified fraud risks. However, we also observe that under-auditing can occur when brainstorming quality is low.

Prior and contemporary studies provide evidence that the extent that auditors benefit from brainstorming depends on the brainstorming method employed (Carpenter 2007; Hunton and Gold 2010). Archival research on fraud has investigated data from multiple public companies, but has not examined auditors' judgments in practice (e.g., Beasley 1996). Our results thus triangulate prior experimental and archival research and extend the fraud literature by providing new and important insights on the costs and benefits of fraud brainstorming for actual audit teams. Further, the descriptive data contained herein further enhance our understanding of the fraud decision-making and brainstorming processes. These data should provide a basis for further modeling of brainstorming and future empirical research in team decision-making.

While it might seem obvious (and arguably required by SAS No. 99) that high-quality brainstorming sessions should exist in practice, our results and recent PCAOB findings indicate substantial variation in brainstorming quality (PCAOB 2007). Indeed, auditors continue to seek better brainstorming methods (Landis et al. 2008). Thus, we contribute to practice by providing auditors and regulators with descriptive data that (1) complement previous PCAOB inspections and (2) highlight best practices for auditors to improve their fraud judgments (e.g., attendance of an IT audit specialist, and occurrence of the session early in planning).

Our study is subject to several limitations. Allowing participants to select their own audits could have biased their responses toward engagements that were memorable, such as when fraud was found. Our approach might have also led participants to select sessions they perceived as high quality or to report in such a way as to bolster their perception of having conducted a high-quality session. Our sample contains a fraud detection rate of 24 percent, which is likely higher than the actual rate of fraud in the population and is higher than the rate reported in earlier studies. Unfortunately, few participants reported details related to the frauds they identified, so we have little understanding of their nature. Sensitivity analyses to control for these engagements suggest that reported results are not driven by engagements on which fraud was detected or fraud risk was high. We also find that, for audits where fraud was detected, participants were not biased to report higher quality brainstorming sessions.

REFERENCES

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

———. 2003. *AICPA Practice Aid Series, Fraud Detection in a GAAS Audit*. SAS No. 99 Implementation Guide. New York, NY: AICPA.

Ali, A., S. Klasa, and O. Z. Li. 2008. Institutional stakeholdings and better-informed traders at earnings announcements. *Journal of Accounting and Economics* 46 (1): 47–61.

Asare, S. K., and A. M. Wright. 2004. The effectiveness of alternative risk assessment and program planning tools in a fraud setting. *Contemporary Accounting Research* 21 (2): 325–352.

Bamber, E. M., D. Snowball, and R. M. Tubbs. 1989. Audit structure and its relation to role conflict and role ambiguity: An empirical investigation. *The Accounting Review* 64 (2): 285–299.

Beasley, M. S. 1996. An empirical analysis of the relation between board of director composition and financial statement fraud. *The Accounting Review* 71 (4): 443–465.

———, J. V. Carcello, and D. R. Hermanson. 1999. *Fraudulent Financial Reporting: 1987–1997 An Analysis of U.S. Public Companies*. New York, NY: Committee of Sponsoring Organizations of the Treadway Commission.

———, and J. G. Jenkins. 2003. A primer for brainstorming fraud risks. *Journal of Accountancy* 196 (6): 32–38.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. New York, NY: John Wiley.

Beneish, M. 1997. Detecting GAAP violations: Implications for assessing earnings management among firms with extreme financial performance. *Journal of Accounting and Public Policy* 16 (3): 271–309.

Bonner, S. E. 1990. Experience effects in auditing: The role of task-specific knowledge. *The Accounting Review* 65 (1): 72–92.

Bradshaw, M. T., B. J. Bushee, and G. S. Miller. 2004. Accounting choice, home bias, and U.S. investment in non-U.S. firms. *Journal of Accounting Research* 42 (5): 795–841.

Brazel, J. F., K. Jones, and M. F. Zimbelman. 2009. Using nonfinancial measures to assess fraud risk. *Journal of Accounting Research* 47 (5): 1135–1166.

Cardinaels, E. 2008. The interplay between cost accounting knowledge and presentation formats in cost-based decision-making. *Accounting, Organizations and Society* 33 (6): 582–602.

Carpenter, T. D. 2007. Audit team brainstorming, fraud risk identification, and fraud risk assessment: Implications of SAS No. 99. *The Accounting Review* 82 (5): 1119–1140.

Chaiken, S., and Y. Trope. 1999. *Dual-Process Theories in Social Psychology*. New York, NY: Guilford Press.

Core, J. E., W. R. Guay, and S. P. Kothari. 2002. The economic dilution of employee stock options: Diluted EPS for valuation and financial reporting. *The Accounting Review* 77 (3): 627–652.

Covrig, V. M., M. L. DeFond, and M. Y. Hung. 2007. Home bias, foreign mutual fund holdings, and the voluntary adoption of international accounting standards. *Journal of Accounting Research* 45 (1): 41–70.

De Dreu, C. K. W. 2003. Time pressure and closing of the mind in negotiation. *Organizational Behavior and Human Decision Processes* 91 (2): 280–295.

——— 2007. Cooperative outcome interdependence, task reflexivity, and team effectiveness: A motivated information processing perspective. *The Journal of Applied Psychology* 92 (3): 628–638.

DeFond, M. L., R. N. Hann, and X. Hu. 2005. Does the market value financial expertise on audit committees of boards of directors? *Journal of Accounting Research* 43 (2): 153–193.

Diehl, M., and W. Stroebe. 1987. Productivity loss in brainstorming groups: Toward the solution of the riddle. *Journal of Personality and Social Psychology* 53 (3): 497–509.

Dowling, C. 2009. Appropriate audit support system use: The influence of auditor, audit team, and firm factors. *The Accounting Review* 84 (3): 771–810.

Echambadi, R., and J. D. Hess. 2007. Mean-centering does not alleviate collinearity problems in moderated multiple regression models. *Marketing Science* 26 (3): 438–445.

Elliott, R. K. 2002. Twenty-first century assurance. *Auditing: A Journal of Practice & Theory* 21 (1): 139–146.

Gibbins, M., S. Salterio, and A. Webb. 2001. Evidence about auditor-client management negotiation concerning client's financial reporting. *Journal of Accounting Research* 39 (3): 535–563.

Glover, S. M., D. F. Prawitt, J. J. Schultz, Jr., and M. F. Zimbelman. 2003. A test of changes in auditors' fraud-related planning judgments since the issuance of SAS No. 82. *Auditing: A Journal of Practice & Theory* 22 (2): 237–251.

Hair, J. F., R. E. Anderson, R. L. Tatham, and W. C. Black. 1998. *Multivariate Data Analysis*. 5th edition. Upper Saddle River, NJ: Prentice Hall.

Hackenbrack, K., and W. R. Knechel. 1997. Resource allocation decisions in audit engagements. *Contemporary Accounting Research* 14 (3): 481–499.

Hill, G. W. 1982. Group versus individual performance: Are N+1 heads better than one? *Psychological Bulletin* 91 (3): 517–539.

Hinsz, V. B., R. S. Tindale, and D. A. Volrath. 1997. The emerging conceptualization of groups as information processors. *Psychological Bulletin* 121 (1): 43–64.

Hoffman, V. B., and M. F. Zimbelman. 2009. Do strategic reasoning and brainstorming help auditors change their standard audit procedures in response to fraud risk? *The Accounting Review* 84 (3): 811–837.

Hunton, J., and A. Gold. 2010. A field experiment comparing the outcomes of three fraud brainstorming procedures: Nominal group, round robin and open discussion. *The Accounting Review* 85 (3): 911–935.

Hyatt, T. A., and D. F. Prawitt. 2001. Does congruence between audit structure and auditors' locus of control affect job performance? *The Accounting Review* 76 (2): 263–274.

Johnstone, K. M., and J. C. Bedard. 2003. Risk management in client acceptance decisions. *The Accounting Review* 78 (4): 1003–1025.

Kennedy, P. 2008. *A Guide to Econometrics*. 6th edition. Cambridge, MA: MIT Press.

Kline, R. B. 2004. *Principles and Practice of Structural Equation Modeling*. 2nd edition. New York, NY: Guilford Publications.

Kruglanski, A. W., and T. Freund. 1983. The freezing and unfreezing of lay inferences: Effects of impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology* 19 (5): 448–468.

Landis, M., S. I. Jerris, and M. Braswell. 2008. Better brainstorming. *Journal of Accountancy* 206 (4): 70–73.

Loebbecke, J. K., M. M. Eining, and J. J. Willingham. 1989. Auditors' experience with material irregularities: Frequency, nature, and detectability. *Auditing: A Journal of Practice & Theory* 9 (1): 1–28.

Lynch, A. L., U. S. Murthy, and T. J. Engle. 2009. Fraud brainstorming using computer-mediated communication: The effects of brainstorming technique and facilitation. *The Accounting Review* 84 (4): 1209–1232.

McVay, S. E. 2006. Earnings management using classification shifting: An examination of core and special items. *The Accounting Review* 81 (3): 501–531.

Nelson, M. W., J. A. Elliott, and R. L. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings management decisions. *The Accounting Review* 77 (Supplement): 175–202.

Neter, J., M. H. Kutner, C. J. Nachtsheim, and W. Wasserman. 1996. *Applied Linear Statistical Models*. 4th edition. New York, NY: McGraw-Hill.

Nieschwietz, R. J., J. J. Schultz, and M. F. Zimbelman. 2000. Empirical research on external auditors' detection of financial statement fraud. *Journal of Accounting Literature* 19: 190–246.

Nunnally, J. D. 1978. *Psychometric Theory*. New York, NY: McGraw-Hill.

O'Keefe, T. B., D. A. Simunic, and M. T. Stein. 1994. The production of audit services: Evidence from a major public accounting firm. *Journal of Accounting Research* 32 (2): 241–261.

Oppenheim, A. N. 1992. *Questionnaire Design, Interviewing, and Attitude Measurement*. New York, NY: Pinter Publishers.

Postmes, T., R. Spears, and S. Cihangir. 2001. Quality of decision making and group norms. *Journal of Personality and Social Psychology* 80 (6): 918–930.

Prawitt, D. F. 1995. Staffing assignments for judgment-oriented audit tasks: The effects of structured audit technology and environment. *The Accounting Review* 70 (3): 443–465.

Public Company Accounting Oversight Board (PCAOB). 2003. Professionalism is primary. Remarks delivered by Douglas R. Carmichael at AICPA National Conference, Washington, D.C., December 12.

———. 2004. PCAOB standing advisory group meeting: Meeting agenda, September 8–9. Available at: http://pcaobus.org/News_and_Events/Events/2004/09-08-09.aspx.

———. 2007. Observations on auditors' implementation of PCAOB standards relating to auditors' responsibilities with respect to fraud. Available at: http://pcaob.org/inspections/other/01-22_release_2007-001.pdf.

———. 2008. Report on the PCAOB's 2004, 2005, 2006, and 2007 inspections of domestic annually inspected firms. Available at: http://pcaobus.org/Inspections/Documents/2008_12-05_Release_2008-008.pdf.

Schippers, M. C., D. N. Den Hartog, P. L. Koopman, and J. A. Wienk. 2003. Diversity and team outcomes: The moderating effects of outcome interdependence and group longevity and the mediating effect of reflexivity. *Journal of Organizational Behavior* 24 (6): 779–802.

Stasser, G. 1999. The uncertain role of unshared information in collective choice. In *Shared Cognition in Organizations: The Management of Knowledge*, edited by Thompson, L., J. M. Levine, and D. M. Messick, 49–69. Hillsdale, NJ: Erlbaum.

———, and Z. Birchmeier. 2003. Group creativity and collective choice. In *Group Creativity: Innovation through Collaboration*, edited by Paulus, P. B., and B. A. Nijstad, 85–109. New York, NY: Oxford University Press.

U.S. House of Representatives. 2002. Sarbanes-Oxley Act of 2002. Public Law 107–204 [H. R. 3763]. Washington, D.C.: Government Printing Office.

Wilks, T. J., and M. F. Zimbelman. 2004. Decomposition of fraud-risk assessments and auditors' sensitivity to fraud cues. *Contemporary Accounting Research* 21 (3): 719–745.

Wooldridge, J. M. 2009. *Introductory Econometrics: A Modern Approach*. 4th edition. Mason, OH: South-Western.

Zimbelman, M. F. 1997. The effects of SAS No. 82 on auditors' attention to fraud risk factors and audit planning decisions. *Journal of Accounting Research* 35 (Supplement): 75–97.