AUDITING: A JOURNAL OF PRACTICE & THEORY
Vol. 28, No. 1
May 2009
pp. 95–111

American Accounting Association
DOI: 10.2308/aud.2009.28.1.95

The Effects of Audit Review Format on Review Team Judgments

Christopher P. Agoglia, Richard C. Hatfield, and Joseph F. Brazel

SUMMARY: The promulgation of standards (e.g., PCAOB 2004b; IFAC 2008b) highlights the importance of workpaper documentation quality and its influence on audit quality. Our study matches audit workpaper preparers with reviewers to examine how alternative workpaper review methods affect sequential audit review team judgments through their impact on preparer workpaper documentation. While reviewers maintain the option of reviewing workpapers on site ("face-to-face review"), they can now also perform their reviews electronically from remote locations ("electronic review") because of technological advancements such as email and electronic workpapers. Recent research has found that review mode can affect the judgments of auditors preparing the workpapers. Our study extends the literature by examining the extent to which review mode (electronic versus face-to-face) affects the quality of documentation in the workpapers and whether reviewers are able to discern and compensate for these documentation quality issues. Our results indicate that reviewers' judgments are ultimately affected by the form of review expected by their preparer. We test two alternative mediation models to provide insight into *why* the review format affects reviewer judgment quality. Mediation analyses suggest that the effect of review mode on reviewer judgments is mediated by a documentation quality assessment gap. Specifically, with electronic review, reviewers' burden to recognize and compensate for lower-quality documentation was generally greater, often resulting in lower-quality reviewer judgments than when the mode of review was face-to-face. These results suggest that the effect of review mode can persist to the reviewer's judgment through its influence on preparer workpaper documentation and the resulting documentation quality assessment gap.

Keywords: documentation; review process; judgment quality; electronic review; face-to-face review.

Christopher P. Agoglia is an Associate Professor at the University of Massachusetts, Richard C. Hatfield is an Associate Professor at The University of Alabama, and Joseph F. Brazel is an Assistant Professor at North Carolina State University.

We thank Brian Daugherty, Michael Favere-Marchesi, Michael Grayson, Diane Matson, Steve Salterio, Ken Trotman, Rick Tubbs, two anonymous reviewers, and workshop participants at the 2005 International Symposium on Audit Research, 2006 AAA Auditing Section Midyear Meeting, and 2006 AAA Annual Meeting for their helpful comments. Professor Agoglia thanks Drexel University for financial support. We appreciate the support of the audit professionals who participated in this study.

Editor's note: Accepted by Ken Trotman.

Submitted: November 2006
Accepted: September 2008
Published Online: May 2009

INTRODUCTION

This study examines how the mode of audit workpaper review affects review team judgments through its influence on the reviewer's ability to appropriately assess and react to quality issues with documentation provided in the workpapers.[1] The technological advancements made in recent years (e.g., electronic workpapers and electronic communication) have provided reviewers with options regarding how they wish to conduct their reviews. Before the advent of electronic workpapers, preparers would pass off hard-copy workpapers to their reviewers. In turn, the reviewers would prepare review notes and, if they were on-site with the preparer, they could choose to discuss these review notes with the preparer in person (face-to-face review).[2] While reviewers maintain the option of a face-to-face interaction with the preparer, the switch to electronic workpapers facilitates remote reviews via electronic communication with the preparer. ''Electronic'' reviews (or e-reviews) help reviewers manage their increasingly heavy workloads by allowing them to review multiple jobs concurrently and by giving them the option of not traveling to a client when the timing and/or location are inconvenient (Agoglia et al. 2008). The widespread use of electronic review has been acknowledged in recent audit guidance, underscoring its increasing significance in the contemporary audit environment (IFAC 2008a).

The review process is designed to serve as a quality control mechanism (AICPA 1978; Rich et al. 1997). The primary objectives of the reviewer are to ensure the adequacy of the preparer's work and to draw his or her own conclusions based on the evidence the preparer documents in the workpapers (Tan 1995; Agoglia et al. 2003). However, the difficulty of the reviewer's task, and in turn the reviewer's effectiveness, may vary depending on the quality of the evidence documented by the preparer (Libby and Trotman 1993). Recent results suggest that alternative methods of review may differentially affect how auditors prepare audit documentation. For example, Brazel et al. (2004) find that preparers regard electronic review as somewhat less demanding than face-to-face review and, consequently, spend less time preparing workpapers, are more likely to anchor on prior-year audit judgments, and are less concerned with pre-review workpaper effectiveness than face-to-face preparers. If the reviewer is unable to identify (and sufficiently mitigate) any resulting workpaper quality issues through review, then the reviewer's choice of review method may eventually impair the review team's judgment. In our study, we consider not only whether the review format anticipated by the preparer affects *reviewer* judgments but, if so, why. We test two models in order to more fully explore the relationship between review format and reviewer judgment quality. One model predicts that preparer workpaper documentation quality mediates this relationship, while the other predicts that this relationship is mediated by the extent to which reviewers appropriately evaluate preparer documentation quality (i.e., actual versus reviewer assessments of documentation quality).

In our study, we matched preparers with reviewers and manipulated review mode in order to examine its effect on review team judgment. Audit seniors (preparers) were provided with prior-year workpapers (which reflected positively on the client's financial condition) and current-year evidence (which indicated declining financial conditions) for a hypothetical client and were asked to document evidence in the current-year workpaper for

[1] Similar to Trotman (1985), we define a review team as consisting of a hierarchical pair of auditors: a subordinate auditor who prepares the workpapers and a supervising auditor who reviews this work, with the review team's efforts culminating in the judgments/decisions of the reviewing auditor.

[2] Face-to-face is still a popular form of review, with the only change being that work is now prepared electronically.

a preliminary going concern conclusion. Prior to the start of their task, preparers were informed that their workpapers would be reviewed, and of the form that the review would take. Half of the preparers anticipated a face-to-face review (i.e., in-person meeting with the reviewer allowing for discussion of review notes), while the other half expected to be reviewed electronically (i.e., review notes would be relayed and discussed via email). We then randomly formed review teams by individually matching audit managers (reviewers) with preparers. Reviewers evaluated their preparer's documentation and provided a going concern assessment. Preparer documentation quality and reviewer judgment quality were assessed with the aid of audit experts.

Results indicate that the effect of review mode persists through the review process and influences the judgments of review teams. Relative to face-to-face review, the quality of reviewers' going concern judgments was significantly lower when their workpaper preparers anticipated an electronic review. Mediation analyses suggest that it is the difference between reviewer perceptions of workpaper documentation quality and actual documentation quality as determined by experts (the ''documentation quality assessment gap''), and not simply the preparer's documentation quality, that mediates the effect of review mode on reviewers' judgment quality. Specifically, we find that documentation quality issues may not be corrected through the review process (i.e., the reviewer may not perceive an issue with documentation quality when one is present) and could ultimately impair the final judgments of reviewers. We speculate as to what may influence the likelihood of reviewer identification of documentation quality issues when they are present.

This study extends the review process literature and informs practice in several ways. First, we consider the effect of review format (a choice typically within the control of the reviewer), which has only been investigated at the preparer level (Brazel et al. 2004; Payne et al. 2007). Our results provide insight into the importance and consequences of this reviewer choice. Second, through our manipulation of the anticipated form of the preparer/reviewer interaction, we focus on sequential audit *review team* judgments. Prior research typically focuses on one of the team members (preparer or reviewer) and considers that individual's reaction to a particular manipulation (e.g., Ricchiute 1999; Tan and Trotman 2003; Brazel et al. 2004; Payne et al. 2007). Third, unlike prior studies, we investigate reviewers' ability to identify and compensate for preparer documentation quality issues. We find that review format influences documentation quality, with a greater burden placed on e-reviewers to identify and compensate for inferior documentation. These reviewers typically did not recognize documentation quality issues and were less likely to plan/take actions to compensate for possible shortcomings in the workpapers (i.e., rely less on workpapers, spend more time on review, request more rework, or provide more review comments to staff). Our results suggest that managers and partners may want to consider the demands and risks of the audit task when choosing their mode of review or take into account (and plan actions to address) the likelihood of differential documentation quality associated with alternative modes of review. Further, firms may want to consider training their reviewers (particularly those who have less experience reviewing electronically) about the pitfalls associated with e-review.

The remainder of this paper is organized as follows. The next section discusses the background and related research and develops hypotheses. This is followed by a discussion of the method and presentation of the results. The final section offers conclusions and implications.

BACKGROUND AND HYPOTHESES DEVELOPMENT

Alternative Methods of Review

Workpaper review is a hierarchical process in which the work of subordinates is scrutinized by more-experienced auditors (e.g., Ashton et al. 1988; Libby and Trotman 1993). A considerable amount of research has examined how the anticipation of workpaper review affects preparers' judgments. Results of this research indicate that, for example, expectation of review can lead to: greater consensus and self-insight, reduction of recency bias, greater documentation of evidence, and higher-quality decisions (e.g., Johnson and Kaplan 1991; Kennedy 1993; Koonce et al. 1995; Tan and Kao 1999). This prior research often focused on comparing the effects on preparers of anticipating an in-person review to no review. The anticipation of an in-person review serves to emulate a traditional form of review where the reviewer and the preparer are in the same location, allowing the reviewer to discuss certain issues or concerns with the preparer in person (*face-to-face review*). In practice, however, electronic workpapers have resulted in the reduced use of face-to-face (on-site) review in recent years, as reviewers are conducting more of their reviews from locations other than the client (Brazel et al. 2004). These *electronic reviews* typically involve the reviewer's examining workpapers online and interacting with the preparer electronically to relay, discuss, and resolve review notes. According to survey evidence, electronic reviews (or reviews with no ''in-person'' interaction) now constitute somewhere between 37 percent and 63 percent of all reviews conducted (Brazel et al. 2004; Fargher et al. 2005). The International Federation of Accountants (IFAC) has recently acknowledged this increased use of electronic reviews and, consequently, has advised that audit planning should explicitly consider whether reviews should occur at the client, off-site, or both (IFAC 2008a).

Previous research suggests that alternative methods of review may result in differing judgments by staff anticipating the review. For example, Brazel et al. (2004) compare the judgments of auditors expecting face-to-face reviews (face-to-face preparers) with those of auditors expecting electronic reviews (e-review preparers) upon completion of their workpapers. They find that, due to differing demands (e.g., perceptions of accountability, synchronicity of communication) perceived by preparers in the two review conditions, face-to-face preparers spend more time preparing a workpaper (i.e., are less efficient) than e-review preparers. Additionally, face-to-face preparers are less likely to be influenced by prior-year judgments (i.e., reduced anchoring), are more concerned with workpaper effectiveness, and provide higher-quality judgments. Payne (2004) finds that, relative to staff auditors who are anticipating a written review, those expecting a face-to-face review devote greater effort, resulting in improved performance of a complex task (identification of a trend in evidence).[3] Payne et al. (2007) examine a somewhat similar review format, referred to as ''review-by-interview,'' which also requires face-to-face interaction, and conclude that anticipation of a review-by-interview leads to greater preparer focus on more cognitively demanding procedures—and, in turn, better preparer performance—than does anticipation of a written review with no face-to-face discussion. Thus, based on the findings of these studies, it appears that preparers perceive reviews involving in-person (on-site) interaction with their reviewer as more demanding and therefore devote more pre-review cognitive effort to the workpapers, likely resulting in higher-quality workpapers. Recent audit guidance highlights the importance of workpaper documentation quality and its significance for those who review the workpapers (PCAOB 2004b; IFAC 2008b).

[3] While Payne (2004) does not investigate electronic review, a written review (as used in her study) is likely quite similar in terms of timing and expected response, with the medium (paper versus electronic file) being the primary difference.

Review Team Judgments

The hierarchical nature of review teams typically results in a sequential process of interactions between reviewer and preparer that centers on the preparer's workpapers (Rich et al. 1997), and culminates in the judgments of the reviewing auditor. This sequential judgment process enhances the quality of audit decisions (e.g., Trotman 1985; Trotman and Yetton 1985; Libby and Trotman 1993). Libby and Trotman (1993) propose that one reason the review process is effective is that preparers and reviewers focus their attention on different types of information (i.e., evidence consistent/inconsistent with the preparer's judgment). However, they also note that the reviewer and preparer do not necessarily evaluate the same evidence set. Reviewers have access to what is recorded in the workpapers by the preparer, and the preparer may have motivation/incentive (e.g., preservation or enhancement of their reputations, improved efficiency/conservation of time and effort) to selectively document evidence more in line with their judgments (Gibbins 1984; Libby and Trotman 1993; Rich et al. 1997). These incentives appear to be influenced by a variety of factors, including the type of review expected by preparers (e.g., Brazel et al. 2004). Thus, although the reviewer will have access to the data set documented by the preparer, this data set may not be as rich as the data set originally observed by the preparer (e.g., omission of key evidence), potentially reducing reviewer effectiveness (Libby and Trotman 1993; Anderson and Koonce 1998). Further, enhanced time constraints due to recent legislation (e.g., PCAOB 2004a; SEC 2005) and advances in technology provide reviewers with the incentive and opportunity to perform review activities off-site, which could potentially increase reviewers' reliance on biased or insufficient preparer documentation.

If the review format affects the quality of the preparer's workpaper documentation, then the choice of review format (a choice typically within the control of the reviewer) likely has significant implications for reviewer judgments. Specifically, relative to face-to-face review, the anticipation of an e-review may lead to less-effective preparer documentation, which, in turn, may result in lower-quality review team judgments if the reviewer fails to recognize and compensate for this lower-quality documentation. However, reviewers may perceive that workpapers will be of lower quality when preparers expect an e-review. Gibbins and Trotman (2002) find that, as reviewer perceptions of the quality of the preparer's workpapers decrease, their effort in the review process increases. Consequently, the review process may mitigate any review mode effects noted at the preparer/pre-review level.

However, prior research suggests that reviewers' judgments are significantly influenced by the evidence/judgments included in the workpapers (Ricchiute 1999; Yip-Ow and Tan 2000; Agoglia et al. 2003). For example, Ricchiute (1997, 1999) finds that partners who are provided evidence consistent with a particular conclusion are more likely to reach that same conclusion relative to partners who receive balanced evidence. Thus, to the extent that the expectation of an electronic review reduces preparers' documentation quality *and* their reviewers are unable to identify (and sufficiently mitigate) these quality issues, these reviewers' judgments will suffer qualitatively relative to the judgments of face-to-face reviewers. This leads to the following expectation.

H1: The quality of reviewers' judgments will be lower when their audit staff prepare workpapers with the expectation of an electronic review versus a face-to-face review.

Examining the Influence of Review Format on Reviewer Judgment

We examine two alternative mediation models to more fully explore the relationship between review format and reviewers' judgment quality. The first considers the link between

preparer documentation quality and reviewer judgment quality. Prior studies suggest that biased preparer documentation typically leads to reviewer judgments that are biased in the same direction (e.g., Ricchiute 1997, 1999). When coupled with other prior research indicating that workpaper quality suffers with electronic review (Brazel et al. 2004), this suggests a link among review format, documentation quality, and reviewers' judgment quality (see Figure 1, Panel A). In other words, preparer documentation quality (resulting from the review mode utilized) may serve as the mechanism influencing reviewer judgments.[4] Therefore, we test the following mediating hypothesis.

H2a: Documentation quality will mediate the effect of review mode on reviewers' judgment quality.

Recall that the development of H1 presupposes that (1) electronic review results in generally lower-quality preparer documentation *and* that (2) these quality issues may be difficult for reviewers to appropriately identify/assess and mitigate. Consistent with this rationale, we would expect there to be a larger *documentation quality assessment gap* (i.e., the difference between reviewer perceptions of workpaper documentation quality and actual documentation quality as determined by experts) for e-reviewers than for face-to-face reviewers. In turn, this larger documentation quality assessment gap likely leads to reviewers being less adequately informed under e-review, resulting in lower-quality reviewer judgments than under face-to-face review. It should be noted that in the studies by Ricchiute (1997, 1999), reviewers had no reason to suspect preparer bias and, thus, may have felt no need to adjust for it. In contrast, reviewers in our study are aware of the review format anticipated by their preparers. It is possible then that, as the reviewer's choice of review method is a controllable/observable attribute of the audit review process, some reviewers may suspect that e-review results in workpapers that are qualitatively different (i.e., of lower quality) than those prepared in expectation of an in-person review (Agoglia et al. 2008). If this is the case, some e-reviewers may take actions that help them increase the likelihood that they identify documentation quality issues and adjust accordingly, thereby improving the quality of their judgments. That is, preparer documentation quality should only affect reviewer judgments when the reviewer is unaware of, and cannot compensate for, preparer documentation quality issues.

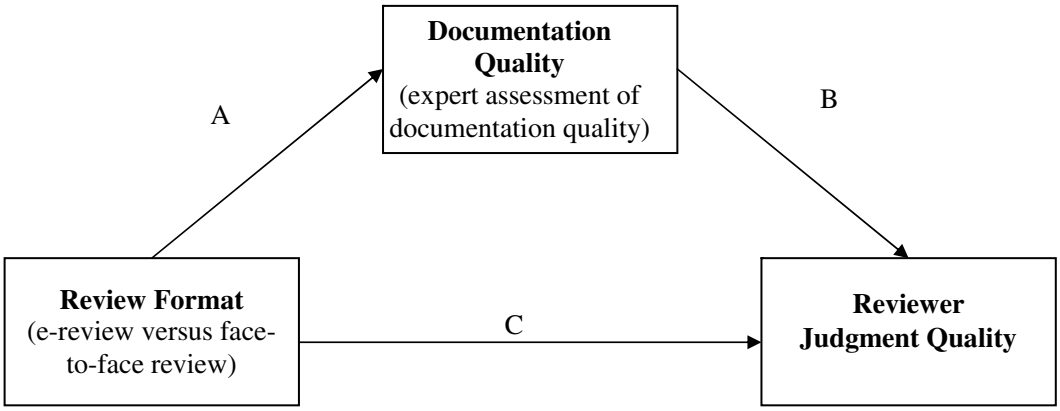
This discussion suggests that the documentation quality assessment gap (as opposed to, simply, the quality of the documentation) may serve as the mechanism behind the effect predicted in H1. Specifically, our setting (in which reviewers are aware of the preparer's expected review mode) allows reviewers to anticipate and potentially mitigate documentation quality effects. Thus, we test a model that predicts that the effect of preparers' anticipated review mode on their reviewers' judgments (H1) will be mediated by the documentation quality assessment gap (see Figure 1, Panel B).

H2b: The reviewers' documentation quality assessment gap will mediate the effect of review mode on reviewers' judgment quality.

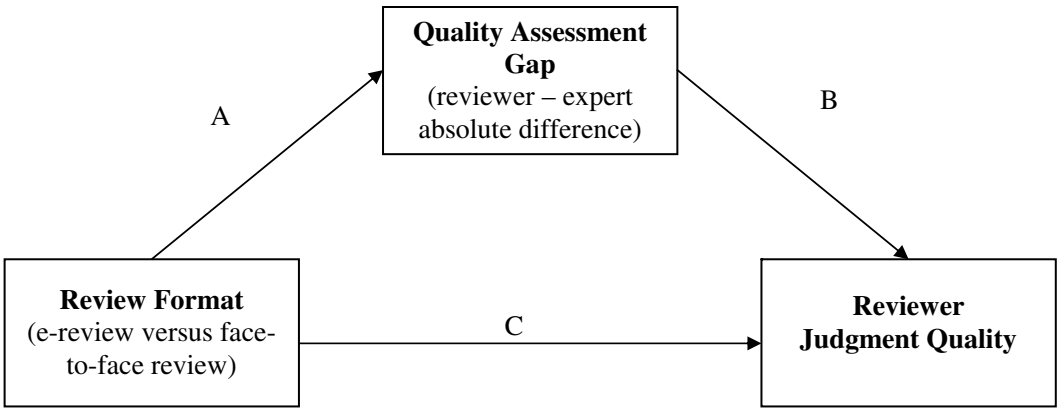
[4] It should be noted that the Ricchiute (1997, 1999) studies examine documentation bias, not quality. The extent to which bias and quality differ may affect whether quality effectively mediates the format/reviewer judgment quality relationship.

FIGURE 1
Alternative Mediation Models

Panel A: Documentation Quality as Mediator (H2a)



Panel B: Quality Assessment Gap as Mediator (H2b)



METHOD

Participants

Sixty practicing auditors from large international public accounting firms participated in this study, 30 as preparers and 30 as reviewers. Auditors participating as preparers were audit seniors with an average of about three years' experience, while auditors participating as reviewers were generally audit managers with an average of more than seven years'

experience.[5] Prior research and discussions with audit managers revealed that auditors with these levels of experience would be familiar with preparing workpapers relating to the financial viability of their clients and with reviewing these workpapers, respectively (e.g., Libby and Trotman 1993; Rau and Moser 1999).

Review Team Experimental Task

Documentation Phase

Preparer participants prepared a preliminary going concern evaluation workpaper for a hypothetical client. The task was adapted from Tan (1995) and completed via computer (see also Brazel et al. 2004). Preparers received detailed instructions regarding the task, relevant authoritative guidance, prior-year workpapers, and current-year audit evidence. Consistent with Tan (1995), prior-year workpapers documented a conclusion that the ''going concern assumption appears reasonable'' and included a corresponding memo summarizing the important evidence.[6] The prior-year memo contained ten items that were largely positive with respect to the financial condition of the client. In contrast, the current-year audit evidence, consisting of 30 items, indicated a decline in financial conditions for the client, with ten items that supported the going concern assumption, ten items that undermined the going concern assumption, and ten irrelevant items (see Tan 1995).[7] Preparers were asked to document (in the current-year workpaper) evidence for a preliminary going concern conclusion.[8]

Preparers were randomly assigned to one of two conditions: face-to-face review or electronic review. Preparers in the face-to-face review group were informed that their workpaper would be reviewed by an audit manager who would later meet with the preparer, allowing for in-person discussion of any related review notes. Electronic review preparers were also advised of the workpaper review, but informed that all correspondence with their reviewers would take place via email.[9,10] All preparers had access to relevant excerpts from Statement on Auditing Standards (SAS) No. 59 (AICPA 1988) regarding the going concern assessment, and viewed identical prior-year going concern evaluation workpapers and

[5] There are no significant differences in audit experience between preparer groups (means = 32.27 and 35.53 months; $t = 0.617$, $p = 0.542$, for the face-to-face and electronic review conditions, respectively) or reviewer groups (means = 7.23 and 7.13 years; $t = 0.083$, $p = 0.935$, for reviewers in the face-to-face and electronic review conditions, respectively). Also, there are no significant differences (all p's > 0.35) between groups on other demographic variables (e.g., perception of the percentage of firms that fail, client size). Data were collected in a period bridging the adoption of PCAOB Auditing Standard No. 3. Preparer data were gathered pre-adoption, while reviewer data and the expert assessment of documentation quality were gathered post-adoption.

[6] As with Tan (1995), the prior-year workpaper also presented the prior-year preparer's preliminary conclusion regarding the reasonableness of the client's going concern assumption numerically (+4 on a scale where −7 = ''definitely not reasonable'' and +7 = ''definitely reasonable'').

[7] Similar to Cohen et al. (2000), we investigate auditor performance under declining financial conditions due to the increased risks auditors face under such conditions (e.g., the failure to modify the audit opinion prior to client bankruptcy).

[8] Preparers documented evidence they considered relevant to the going concern assessment from the available evidence. Preparers were also asked to provide a preliminary assessment of the reasonableness of the client's going concern assumption on the same −7 to +7 scale as presented in the prior-year workpaper.

[9] In practice (and in our study), workpapers and related review notes are generally prepared and submitted electronically. We manipulate whether the reviewer will be present to potentially discuss certain issues relating to written review notes. These conditions represent the two most common forms of review in practice (Brazel et al. 2004; Fargher et al. 2005).

[10] A manipulation check revealed that preparers understood and anticipated their respective review conditions (face-to-face or electronic review). One participant in the face-to-face review condition indicated that he did not expect a face-to-face review. Removing this participant's responses from the analysis does not affect the conclusions drawn. Preparers later met or corresponded via email with a reviewer. Logistical constraints necessitated that these reviewers (nine, in total, who corresponded with all 30 preparers) were individuals other than the 30 reviewer participants in our study (who each reviewed only a single preparer's documentation).

current-year audit facts (which were randomized to control for order effects). Preparers were able to access this data while preparing their workpapers. They then answered a series of demographic and case-related questions.

Review Phase

Reviewer participants were randomly matched with a preparer and asked to assume the role of audit manager on the preparer's engagement. Reviewers were provided with relevant audit guidance, the prior-year going concern evaluation workpaper, and the preparer's current-year going concern workpaper. Reviewers were instructed to review the preparer's workpaper and provide review notes to the preparer. Reviewers were told to assume they would be either meeting in person with their preparer (allowing for discussion of reviewer concerns) or communicating electronically with their preparer, depending on (corresponding to) their preparer's condition (i.e., reviewers were aware of their preparer's review condition). Reviewers prepared their review notes in an electronic document. After completing their reviews, they recorded the time it took them to perform their reviews. On separate 11-point scales, reviewers also assessed: the ''quality of the supporting documentation'' provided by their preparer (0 = ''very low''; 10 = ''very high''); ''how much [they] would rely on the going concern documentation'' provided by their preparer (0 = ''no reliance''; 10 = ''complete reliance''); and ''how much additional work'' they expected the preparer would ''need to do to produce the final version of this workpaper'' (0 = ''no rework''; 10 = ''substantial rework''). In addition, reviewer participants provided their own assessment of the reasonableness of the going concern assumption for the hypothetical client on a 15-point scale, with end points labeled ''definitely not reasonable'' (coded as −7) and ''definitely reasonable'' (coded as +7). They then answered a series of demographic and case-related questions.

Measuring Judgment Quality

In this study, we examine the quality of reviewers' judgments, as opposed to solely the direction or magnitude of their judgments. To determine the quality of reviewers' judgments, we enlisted the aid of three experts (all audit partners with an average of 10.6 years' experience) from large international firms. Our experts were supplied with the same materials as preparers and asked to provide a going concern assessment based on all the available evidence (using the same 15-point scale as reviewers). Following Tan (1995), we measure judgment quality by computing the absolute deviations of reviewers' assessments from the mean of our experts. After providing their going concern assessments, experts were also asked to individually evaluate ''the quality of the supporting documentation provided'' by each of the 30 preparers on an 11-point scale from 0 (''very low'') to 10 (''very high''). This is the same scale reviewers used to record their assessments of preparer documentation quality. These measures were used to determine each reviewer's ''documentation quality assessment gap,'' computed as the absolute difference between the reviewer's assessment of preparer documentation quality and the experts' mean assessment of preparer documentation quality. Experts were unaware of preparer review conditions.

RESULTS

Review Team Judgment Quality (H1)

Hypothesis 1 predicts that reviewer judgments will be of higher quality (i.e., lower deviation from the expert mean) when the workpapers are prepared under the anticipation of a face-to-face review than an electronic review. Nontabulated results indicate that the mean assessment of reviewers of face-to-face preparers (0.67) mirrors that of the experts

(mean = 0.67).[11] In contrast, the mean assessment of reviewers of electronic review preparers (2.33) differs considerably from that of the experts. Mean absolute differences from the experts' assessment are significantly smaller for the face-to-face review group (1.38) than for the electronic review group (2.73; p = 0.003; Table 1), providing support for H1.[12] Thus, although the review process is designed to serve as a quality control mechanism, the choice of review mode may affect its ability to serve that function and ultimately may affect the quality of review team judgments.

Mediation Analysis (H2)

Recall that our expectation of lower-quality judgments for electronic review teams is based on the premise that: a) e-review results in generally lower-quality preparer documentation *and* b) these quality issues may be difficult for reviewers to appropriately identify. We find support for the former notion through our experts' assessments of preparer documentation quality. Results indicate that the experts consider the evidence documentation provided by the face-to-face group to be of higher quality than that of the e-review group (mean expert quality rating = 5.69 versus 3.79, respectively; p = 0.001; Table 1).[13] Further, consistent with the latter notion, results suggest that some reviewers have trouble recognizing these e-review documentation quality issues. Specifically, the documentation quality assessment gap (i.e., reviewer-expert absolute differences in assessment of preparer documentation quality) is significantly greater for the electronic review condition than for the face-to-face condition (mean absolute reviewer-expert differences = 2.63 and 1.29, respectively; p = 0.004; Table 1). An examination of the mean responses from reviewers and experts indicates that, relative to the experts, reviewers perceived the e-review preparers' supporting documentation to be of significantly higher quality (means = 6.53 and 3.79 for the reviewers and experts, respectively; p < 0.001, two-tailed). No such significant difference occurred with the higher-quality face-to-face preparers' documentation (means = 5.33 and 5.69 for the reviewers and experts, respectively; p = 0.461, two-tailed).[14]

We explicitly consider the role preparers' workpaper documentation plays in their reviewers' judgments. Specifically, H2a tests whether documentation quality mediates the relationship between review format and reviewer judgment quality, while H2b examines the mediating influence of reviewers' ability to appropriately assess the quality of their preparers' documentation (i.e., the documentation quality assessment gap). In order to test these mediation hypotheses, we follow the procedures outlined in Baron and Kenny (1986) and estimate the following regressions:

---

[11] The mean going concern assessment of the experts in our study (0.67) is consistent with that obtained from four audit partners (0.50) who were provided with identical prior-year and current-year information in Tan (1995).

[12] Comparisons between face-to-face and electronic review conditions are tested using t-tests. Due to the directional nature of expectations, all tests are one-tailed unless otherwise stated.

[13] In addition, experts were asked to provide a list (based on the full set of current-year evidence) of the ''most important of these evidence items with respect to the data that [they] feel would be most essential for an audit senior to bring to the attention of a reviewer (audit manager) when evaluating the client's ability to continue as a going concern.'' A total of 12 items were listed as ''most important'' by the experts, with face-to-face preparers documenting a significantly greater number of these ''important items'' in their workpapers (mean = 6.07 items) than e-review preparers (mean = 4.33 items; p = 0.002; nontabulated).

[14] Recall that, while reviewers had the current-year workpaper (i.e., the current-year audit evidence items *documented* in the workpaper by their preparers), the experts had access to the evidence underlying the current-year workpaper (i.e., all current-year evidence *observed* by the preparers). Both the experts and the reviewers had access to the prior-year workpapers, as well as relevant authoritative guidance.

TABLE 1
Workpaper Documentation Quality and Reviewer Judgment Quality

Variable[a]		Face-to-Face Review	Electronic Review	t-statistic[b]	p-value
Reviewer Judgment Quality (reviewer-expert absolute difference) **(H1)**	Mean	1.38	2.73	3.03	0.003
	(SD)	1.28	1.17		
Expert Assessment of Documentation Quality	Mean	5.69	3.79	3.42	0.001
	(SD)	1.62	1.42		
Reviewer Assessment of Documentation Quality	Mean	5.33	6.53	1.44	0.160
	(SD)	2.61	1.89		
Documentation Quality Assessment Gap (reviewer − expert absolute difference)	Mean	1.29	2.63	2.92	0.004
	(SD)	1.13	1.38		

[a] *Reviewer Judgment Quality* was computed as the absolute deviation of each reviewer's going concern assessment from the experts' mean going-concern assessment. Reviewer participants (as well as experts) were prompted to "provide your own going-concern assessment" for the hypothetical client on a 15-point scale from −7 ("definitely not reasonable") to +7 ("definitely reasonable"). The *Expert* and *Reviewer Assessment of Documentation Quality* represent the mean responses of the experts and reviewers, respectively, to the prompt "The quality of the supporting documentation provided by this audit senior can be described as." Responses were recorded on an 11-point scale where 0 = "very low" and 10 = "very high." The *Documentation Quality Assessment Gap* was computed as the absolute deviation of each reviewer's assessment of documentation quality from the experts' mean assessment of documentation quality.

[b] All tests are one-tailed due to the directional nature of expectations, except for the *Reviewer Assessment of Documentation Quality*. For this measure, no directional expectations were made and therefore a two-tailed test was performed.

$$MED_i = \alpha_0 + \alpha_1 R_i + \varepsilon_i; \quad (1)$$

$$RJQ_i = \gamma_0 + \gamma_1 RF_i + \varepsilon_i; \quad (2)$$

$$RJQ_i = \delta_0 + \delta_1 RF_i + \delta_2 MED_i + \varepsilon_i. \quad (3)$$

MED refers to the mediator in the model, either the experts' assessment of documentation quality (*DQ*) to test H2a or the reviewer's documentation quality assessment gap calculated as the absolute difference between the reviewer's assessment of documentation quality and the mean of the three experts' assessment of documentation quality (*QAG*) for H2b. *RF* refers to the review format and is operationalized as 0 if the workpaper was prepared by staff expecting a face-to-face review and 1 if the workpaper was prepared by staff expecting an electronic review. *RJQ* represents the quality of a reviewer's going concern judgment and is calculated as the absolute difference between the reviewer's judgment and the mean judgment of the three audit experts.

According to Baron and Kenny (1986), the mediating effects predicted by Hypothesis Set 2 would be indicated if the following three conditions were met. First, the independent variable (*RF*) must affect the presumed mediator (*MED*). This is path A in Figure 1 and α_1 in Equation (1) above. Second, the independent variable must affect the dependent variable (*RJQ*). This is represented by path C in Figure 1 and γ_1 in Equation (2) above. Third, the mediator must affect the dependent variable in the presence of the independent variable. This is path B in Figure 1 and δ_2 in Equation (3) above. If all of these conditions hold, then the effect of the independent variable on the dependent variable must be less in Equation (3) than in Equation (2), indicating mediation.

Table 2 reports regression results for Equations (1) through (3). To test for the mediation effects predicted by Hypothesis Set 2, we focus on whether the three conditions discussed above have been satisfied.[15] With respect to H2a, results do not support a mediating role for documentation quality. While Regressions (1) and (2) are consistent with the expectations of H2a, when documentation quality is introduced to the model (Regression (3)), the coefficient on *RF* does not decrease (from 1.355 to 1.400), and *DQ* is not significant ($p = 0.880$). Therefore, documentation quality does not mediate the effect of review format on reviewer judgment quality.

Next we consider H2b and the mediating influence of reviewers' ability to appropriately assess the quality of their preparers' documentation (the documentation quality assessment gap). Table 2 and Figure 2 report regression results for Equations (1) through (3), with the

TABLE 2
Alternative Mediation Models

Panel A: Regression Results for Test of *DQ* as Mediator[a] (H2a)

	Coefficients		
Dependent Variable	**Intercept**	**Review Format (*RF*)**	***DQ***
Documentation Quality ***(DQ)***	7.589	−1.900	
Standard error	.878	.555	
p-value (two-tailed)	.001	.002	
Reviewer Judgment Quality ***(RJQ)***	1.377	1.355	
Standard error	.316	.447	
p-value (two-tailed)	.001	.005	
Reviewer Judgment Quality ***(RJQ)***	−.159	1.400	.024
Standard error	1.380	.542	.155
p-value (two-tailed)	.909	.016	.880

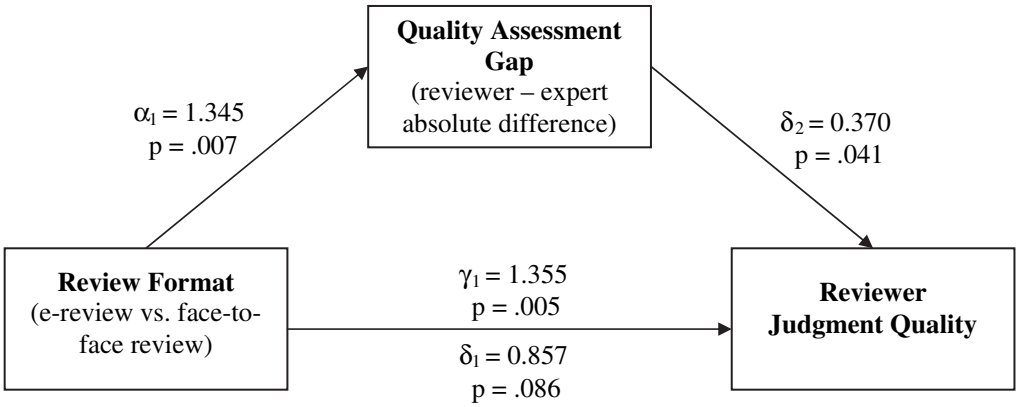
Panel B: Regression Results for Test of *QAG* as Mediator[a] (H2b)

	Coefficients		
Dependent Variable	**Intercept**	**Review Format (*RF*)**	***QAG***
Quality Assessment Gap ***(QAG)***	1.289	1.345	
Standard error	.326	.461	
p-value (two-tailed)	.001	.007	
Reviewer Judgment Quality ***(RJQ)***	1.377	1.355	
Standard error	.316	.447	
p-value (two-tailed)	.001	.005	
Reviewer Judgment Quality ***(RJQ)***	.900	.857	.370
Standard error	.372	.481	.173
p-value (two-tailed)	.023	.086	.041

[a] *Review Format* (RF) was coded as 0 for face-to-face review and 1 for electronic review; *Documentation Quality* (*DQ*) is the mean of the experts' assessments of preparer documentation quality; *Quality Assessment Gap* (*QAG*) was computed as the absolute difference between reviewer and expert assessments of documentation quality; and *Reviewer Judgment Quality* (RJQ) represents the absolute difference between reviewer and expert going-concern assessments.

[15] For the mediation analyses, all reported p-values are two-tailed.

FIGURE 2
Mediation Model-Path Coefficients
Quality Assessment Gap as Mediator (H2b)



The coefficient for the direct path from *Review Format* to *Reviewer Judgment Quality* is represented by γ_1 when the mediator (*Quality Assessment Gap*) is not included in the regression and δ_1 when the mediator is included in the regression. The significance level for γ_1 is .005 increasing to .086 for δ_1. The coefficients for the indirect paths (α_1 and δ_2) are both significant ($p < .05$) See Table 2.

The regressions are specified as follows:

$$QAG_i = \alpha_0 + \alpha_1 RF_i + \varepsilon_i \quad (1)$$

$$RJQ_i = \gamma_0 + \gamma_1 RF_i + \varepsilon_i \quad (2)$$

$$RJQ_i = \delta_0 + \delta_1 RF_i + \delta_2 QAG_i + \varepsilon_i \quad (3)$$

To determine the significance of the mediating effects, the Z-statistic is calculated (MacKinnon and Dwyer 1993) with the coefficients and standard deviations (σ) from Equations (1) and (3) using the following formula:

$$Z = \frac{\alpha_1 \delta_2}{\sqrt{\alpha_1^2 \sigma_\delta^2 + \delta_2^2 \sigma_\alpha^2 - \sigma_\alpha^2 \sigma_\delta^2}}.$$

documentation quality assessment gap (*QAG*) serving as the mediator (*MED*) in these equations. The results of estimating Equation (1) reveal that the coefficient on *RF* is positive and significant ($p = 0.007$), satisfying condition one. Similarly, the results of estimating Equation (2) reveal that the coefficient on *RF* is positive and significant ($p = 0.005$), satisfying condition two. The results of estimating Equation (3) reveal that the coefficient on *QAG* is positive and significant ($p = 0.041$) and that the influence of *RF* is reduced (from $p = 0.005$ to $p = 0.086$), satisfying condition three. We further find that the mediation effect predicted by H2b is significant ($Z = 1.79$, $p = 0.037$; see Figure 2). Thus, our results demonstrate that the documentation quality assessment gap mediates the effect of review format on reviewer judgment quality.

This mediating model (Figure 2) provides a unique lens through which we can view the relationship among review format, preparer documentation, and reviewer judgments, and builds on the contributions made by prior studies (e.g., Ricchiute 1999; Brazel et al. 2004). For example, where Ricchiute (1999) finds that preparers tend to bias documentation toward their conclusions and, in a separate experiment, that biased documentation can lead to biased reviewer judgments, we find that poor-quality documentation (alone) does not necessarily lead to poor reviewer judgments (i.e., H2a is not supported). In the Ricchiute study, however, reviewers had no reason to suspect preparer bias and, thus, may have felt no need to adjust for it. In contrast, our reviewers are aware of the review mode and, given that Agoglia et al. (2008) find that reviewers believe that e-reviews are less effective, may have reason to suspect that workpapers created under e-review may be qualitatively different from those prepared under face-to-face review. In fact, prior research in tax settings suggests that reviewers tend to discount preparer-documented evidence and judgments when they suspect issues with this documentation (Hatfield 2001; Barrick et al. 2004).[16]

Potential Reviewer Actions to Compensate for Documentation Quality Issues

Our results suggest that, in general, the effect of review mode persists to reviewers' judgments through its influence on workpaper documentation and reviewers' inability to detect this influence (i.e., the documentation quality assessment gap). However, in practice, reviewers may choose to: rely less on preparers' initial documentation, increase the amount of rework they require their preparers to perform, provide more review notes to their preparers, and/or spend more time reviewing the workpapers. While any of these actions could potentially help reduce the effect of review mode, additional data gathered suggests that, in general, reviewers of e-review preparers were no more likely to utilize these actions than reviewers of face-to-face preparers. Reviewers were asked to assess (on 11-point scales) how much they felt they could rely on their preparers' going concern documentation (0 = ''no reliance''; 10 = ''complete reliance'') and how much additional work they expected the preparer would ''need to do to produce the final version of this workpaper'' (0 = ''no rework''; 10 = ''substantial rework''). These differences are not significant for reliance on their preparers' documentation (nontabulated means = 4.47 and 3.80, respectively, for e-review and face-to-face groups; $p = 0.307$, two-tailed) and for rework (nontabulated means = 5.20 and 6.33, respectively, for e-review and face-to-face groups; $p = 0.145$, two-tailed). Similarly, reviewers of e-review preparers neither provided significantly more review notes to their preparers (nontabulated means = 6.40 and 6.13, respectively, for e-review and face-to-face groups; $p = 0.664$, two-tailed) nor took significantly longer to perform their reviews than reviewers of face-to-face preparers (nontabulated means = 18.8 and 17.5 minutes, respectively, for e-review and face-to-face groups, $p = 0.456$, two-tailed). Thus, our e-reviewers, in general, were unlikely to take additional measures to compensate for/respond to the typically lower-quality documentation of their preparers' workpapers.[17]

[16] One question raised by our mediation results is ''Why do some e-reviewers recognize and compensate for poor-quality documentation while others do not?'' One study suggests that reviewer task-specific experience (e.g., experience e-reviewing) may reduce the likelihood that errors will flow through the review process (Agoglia et al. 2009). An exploratory analysis dichotomizing the full sample at the median QAG suggests that this may be the case. However, the limitations of our data allow only for speculation, due to issues of small sample size resulting from the median split, and thus we do not provide a detailed description of these results here.

[17] Again, exploratory analysis (see footnote 16) suggests that e-reviewers who were able to recognize quality issues (i.e., more-experienced e-reviewers) were also more likely to take actions to help reduce the effect of review mode on their preparers' documentation.

DISCUSSION AND CONCLUDING REMARKS

Technological enhancements, which have made the use of electronic workpapers and electronic communication commonplace in public accounting, provide reviewers with options regarding how they wish to conduct their reviews and change the structure of the review process. We examine the two most common forms of review in practice: face-to-face review (in which reviewers write review notes and are available to discuss certain items with their preparers in person) and e-review (in which reviewers write review notes and correspond electronically with preparers to resolve them) (Brazel et al. 2004; Fargher et al. 2005). Our study extends the literature by demonstrating that the form of review used by reviewers can ultimately affect their own judgments through its influence on preparer workpaper documentation and the resulting documentation quality assessment gap (i.e., actual versus reviewer assessments of documentation quality).

Results indicate that e-review preparers provided their reviewers with lower-quality workpaper documentation than face-to-face preparers. More critically, the effect of review mode appears to persist through the review process. As it is often difficult for reviewers to identify when the documentation provided by their preparers is of poor quality, e-reviewers' judgments tended to be unduly affected by their preparer's less-informative documentation. Specifically, reviewers' going concern judgments were of lower quality when their workpaper preparers anticipated an electronic review as opposed to a face-to-face review. Further, even though they may perceive e-review to be less effective in general (Agoglia et al. 2008), e-reviewers typically did not recognize or take additional measures to compensate for the lower quality of their preparers' workpapers (i.e., on average, e-reviewers did not: place less reliance on their preparers' documentation, spend more time reviewing, ask for more rework, or provide more comments than face-to-face reviewers).

To gain insight into the mechanism driving this effect, we test two alternative mediation models in order to explore the relationship between review format and reviewer judgment quality. One model predicts that preparer workpaper documentation quality mediates this relationship, while the other predicts that this relationship is mediated by the extent to which reviewers appropriately evaluate preparer documentation quality (i.e., actual versus reviewer assessments of documentation quality). Mediation analyses suggest that e-reviewers' difficulty identifying and mitigating lower-quality workpaper documentation (i.e., the ''documentation quality assessment gap'') often resulted in the reduced quality of their going concern judgments.

The findings of this study have implications for future research. Our study does not investigate specific aspects of documentation quality. Future studies could probe deeper into the notion of documentation quality by examining issues such as documentation completeness and clarity and the resulting effects on review team judgments. Also, as our study involves only two review formats and a single task/context, we cannot speak to the effects of other review formats in other contexts and tasks. Further work could, for instance, investigate more-rudimentary tasks, reviewed using these and other modes of review, to help firms identify which review mode provides the best balance of effectiveness and efficiency for a given task. Future research could also investigate the effect on the documentation quality assessment gap when the preparer is unaware of the mode of review to be applied. In addition, future studies could more thoroughly examine the effect of reviewers' experience with a specific review mode on their evaluation of the evidence their preparer's document in the workpapers, as well as on the resulting review team judgment. Such research will further our understanding of how alternative review formats affect audit quality.

REFERENCES

Agoglia, C. P., T. Kida, and D. M. Hanno. 2003. The effects of alternative justification memos on the judgments of audit reviewees and reviewers. *Journal of Accounting Research* 41 (March): 33–46.

———, J. F. Brazel, R. C. Hatfield, and S. B. Jackson. 2008. The effect of risk of misstatement and workload pressure on the choice of workpaper review format. Working paper, University of Massachusetts.

———, C. A. Beaudoin, and G. T. Tsakumis. 2009. The effect of documentation structure and task-specific experience on auditors' ability to identify control weaknesses. *Behavioral Research in Accounting* 21 (1): 1–18.

American Institute of Certified Public Accountants (AICPA). 1978. *Planning and Supervision*. Statement on Auditing Standards No. 22. New York, NY: AICPA.

———. 1988. *The Auditor's Consideration of an Entity's Ability to Continue as a Going Concern*. Statement on Auditing Standards No. 22. New York, NY: AICPA.

Anderson, U., and L. Koonce. 1998. Evaluating the sufficiency of causes in audit analytical procedures. *Auditing: A Journal of Practice & Theory* 17 (1): 1–12.

Ashton, R. A., D. N. Kleinmuntz, J. B. Sullivan, and L. A. Tomassini. 1988. Audit Decision Making. In *Research Opportunities in Auditing: The Second Decade*, edited by A. R. Abdel-khalik and I. Solomon. Sarasota, FL: American Accounting Association.

Baron, R. M., and D. A. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51 (6): 1173–1182.

Barrick, J. A., C. B. Cloyd, and B. C. Spilker. 2004. The influence of biased tax research memoranda on supervisors' initial judgments in the review process. *The Journal of the American Taxation Association* 26 (Spring): 1–20.

Brazel, J. F., C. P. Agoglia, and R. C. Hatfield. 2004. Electronic versus face-to-face review: the effects of alternative forms of review on auditors' performance. *The Accounting Review* 76 (4): 949–966.

Cohen, J. R., G. Krishnamoorthy, and A. M. Wright. 2000. Evidence on the effect of financial and nonfinancial trends. *Auditing: A Journal of Practice & Theory* 19 (Spring): 27–48.

Fargher, N. L., D. Mayorga, and K. T. Trotman. 2005. A field-based analysis of workpaper review. *Auditing: A Journal of Practice & Theory* 24 (November): 85–110.

Gibbins, M. 1984. Propositions about the psychology of professional judgment in public accounting. *Journal of Accounting Research* 22: 103–125.

———, and K. Trotman. 2002. Audit review: Managers' interpersonal expectations and conduct of the review. *Contemporary Accounting Research* 19 (Fall): 411–444.

Hatfield, R. C. 2001. The effect of staff accountant objectivity in the review and decision process: A tax setting. *Journal of the American Taxation Association* 23 (Spring): 61–74.

International Federation of Accountants (IFAC). 2008a. *Planning an Audit of Financial Statements*. International Standard on Auditing (ISA) No. 300. New York, NY: IFAC.

———. 2008b. *Audit Documentation*. International Standard on Auditing (ISA) No. 230. New York, NY: IFAC.

Johnson, V. E., and S. E. Kaplan. 1991. Experimental evidence on the effects of accountability on auditor judgments. *Auditing: A Journal of Practice & Theory* 10 (Supplement): 96–107.

Kennedy, J. 1993. Debiasing audit judgment with accountability: A framework and experimental results. *Journal of Accounting Research* 31 (Autumn): 231–245.

Koonce, L., U. Anderson, and G. Marchant. 1995. Justification of decisions in auditing. *Journal of Accounting Research* 33 (Autumn): 369–384.

Libby, R., and K. T. Trotman. 1993. The review process as a control for differential recall of evidence in auditor judgments. *Accounting, Organizations and Society* 18 (August): 559–574.

MacKinnon, D., and J. Dwyer. 1993. Estimating mediated effects in prevention studies. *Evaluation Review* 17 (April): 144–158.

Payne, E. A. 2004. An examination of audit workpaper review methods and other factors affecting reviewee performance. Working paper, University of Louisville.

———, R. J. Ramsay, and E. M. Bamber. 2007. The effect of real-time review by interview on auditors' procedures and effectiveness. Working paper, University of Louisville.

Public Company Accounting Oversight Board (PCAOB). 2004a. *An Audit of Internal Control over Financial Reporting in Conjunction with an Audit of Financial Statements*. Auditing Standard No. 2. Washington, D.C.: PCAOB.

———. 2004b. *Audit Documentation.* Auditing Standard No. 3. Washington, D.C.: PCAOB.

Rau, S. E., and D. V. Moser. 1999. Does performing other audit tasks affect going-concern judgments? *The Accounting Review* 74 (October): 493–508.

Ricchiute, D. N. 1997. Effects of judgment on memory: Experiments in recognition bias and process dissociation in a professional judgment task. *Organizational Behavior and Human Decision Processes* 70 (1): 27–39.

———. 1999. The effect of audit seniors' decisions on working paper documentation and on partners' decisions. *Accounting, Organizations and Society* 24 (April): 155–171.

Rich, J. S., I. Solomon, and K. T. Trotman. 1997. The audit review process: A characterization from the persuasion perspective. *Accounting, Organizations and Society* 22 (July): 481–505.

Securities and Exchange Commission (SEC). 2005. *Revisions to Accelerated Deadlines for Filing Periodic Reports*. Release No. 33–8644. December 21. Washington, D.C.: SEC.

Tan, H. 1995. Effects of expectations, prior involvement, and review awareness on memory for audit evidence and judgment. *Journal of Accounting Research* 37 (Spring): 113–135.

———, and A. Kao. 1999. Accountability effects on auditors' performance: The influence of knowledge, problem-solving ability, and task complexity. *Journal of Accounting Research* 37 (Spring): 209–223.

———, and K. T. Trotman. 2003. Reviewers' responses to anticipated stylization attempts by preparers of audit workpapers. *The Accounting Review* 78 (2): 581–605.

Trotman, K. T. 1985. The review process and the accuracy of auditor judgments. *Journal of Accounting Research* 23 (Autumn): 740–52.

———, and P. W. Yetton. 1985. The effect of the review process on auditor judgments. *Journal of Accounting Research* (Spring): 256–67.

Yip-Ow, J., and H. T. Tan. 2000. Effects of the preparer's justification on the reviewer's hypothesis generation and judgment in analytical procedures. *Accounting, Organizations and Society* 25 (February): 203–15.