Electronic vs. Face-to-Face Review: The Effects of Alternative Forms of Review on Auditors’ Performance

Joseph F. Brazel
Department of Accounting
College of Management
North Carolina State University

Christopher P. Agoglia
Department of Accounting
Bennett S. LeBow College of Business
Drexel University
3141 Chestnut Street
Philadelphia, Pennsylvania 19104
Telephone: 215 895-6007
Fax: 215 895-6279
e-mail: cpa22@drexel.edu

Richard C. Hatfield
Department of Accounting
College of Business
University of Texas at San Antonio

We would like to thank Tom Kida, Mark Peecher, Jay Rich, Jean Bedard, Hank Jaenicke, Jim Bierstaker, Terence Ng, Hun Tong Tan, Scott Jackson, participants at the 2002 International Symposium on Audit Research and the 2002 AAA Annual Meeting, and workshop participants at the University of Mississippi, Mississippi State University, University of Vermont, Kansas State University, University of Montana, and North Carolina State University for their helpful comments. We appreciate the assistance of the partner, administrative staff, and audit professionals of the firm that participated in our study. Data is available upon request.
Electronic vs. Face-to-Face Review: The Effects of Alternative Forms of Review on Auditors’ Performance

ABSTRACT: Due to recent technological advancements such as online workpapers and e-mail, audit firms have alternative methods of workpaper review that they did not have in the past. While audit workpaper preparers typically know they will be reviewed, and the form their review will take, prior research has focused on comparing the judgments of auditors who expect to be reviewed with auditors who expect to remain anonymous. This study examines the effects on preparers of using two different methods of review: face-to-face and electronic review. The study also compares both review groups to a no-review control group. Consistent with the Heuristic-Systematic Model, we find that the method of review affects preparer effectiveness and efficiency. Specifically, preparers anticipating a face-to-face review were more concerned with audit effectiveness, produced higher quality judgments, were less efficient at their task, were less likely to be influenced by prior year workpapers, and felt more accountable than preparers in both the electronic review and no-review conditions. Interestingly, electronic review preparers generally did not differ from the no-review group. These results may suggest that how a review is to be conducted, and not merely the expectation that a review will occur, affects the decision maker’s judgments and perceptions.

Keywords: review process; audit effectiveness; audit efficiency; accountability

Data Availability: Data is available upon request.
I. INTRODUCTION

This study examines the effects of alternative methods of audit workpaper review on the performance of auditors who prepare the workpapers. The audit review process is an essential part of an audit (Trotman 1985; Libby and Trotman 1991; Tan 1995). However, it is a costly process, with more than 50 percent of audit manager time and 30 percent of total audit hours allocated to review (Bamber and Bylinski 1987; Asare and McDaniel 1996). Given the significant allocation of auditor time and audit costs to the review process, firms are carefully examining workpaper review in their efforts to re-engineer their audit processes (Rich et al. 1997). In addition, with the technological advancements of the past decade, firms have at their disposal alternative methods of review that they did not have in the past. For example, in contrast to a traditional method of review in which a reviewer and the auditor who prepared the workpapers (preparer) meet face-to-face to discuss the reviewer’s concerns, today online workpapers can be e-mailed from a preparer to a reviewer. The reviewer can review the workpapers online and return them with related review notes via e-mail.

Given that feedback through the review process may take different forms, preparers may view these forms as relatively more or less demanding. These varying demands of different review methods may affect preparer performance. While prior research indicates that the anticipation of a review can, under certain circumstances, affect preparer judgments (e.g., Johnson and Kaplan 1991; Lord 1992; Kennedy 1993; Tan 1995; and Tan and Kao 1999), these studies have typically focused on differences in judgment between auditors who expected their work to be reviewed and auditors who expected their work to remain anonymous. In practice, however, preparers expect that their work will be reviewed (Johnson and Kaplan 1991; Ismail and Trotman 1995). Furthermore, they are typically aware of how that review will be
conducted. Thus, while Tan and Kao (1999) suggest that varying the demands of the review process may be one way firms can influence auditor performance, little is known about the potential effects that alternative review methods may have on preparers.

In this study, two review formats currently used in practice were investigated: (1) face-to-face review in which preparers meet with reviewers and discuss review notes in person and (2) electronic review in which review notes are e-mailed to preparers, allowing preparers more time to formulate their initial responses. Audit seniors were given prior year workpapers (which reflected positively on the client’s financial condition) and current year evidence (which indicated declining financial conditions) and were asked to make and document a preliminary going concern assessment. These auditors were informed that their work would be reviewed, and of the format that review would take (either face-to-face or via e-mail). A third group of audit seniors were told their work would remain confidential. Preparer effectiveness was determined with the aid of audit experts.

The results of this study are consistent with the Heuristic-Systematic Model which predicts, in this context, greater emphasis on effectiveness (relative to efficiency) for face-to-face preparers compared to electronic review preparers. Results indicate that face-to-face preparers were more concerned with the pre-review effectiveness of their going-concern workpapers and arrived at higher quality pre-review judgments, but took longer to prepare their workpapers than preparers in the electronic review group. Also, given the client’s declining financial condition, face-to-face preparers appeared to appropriately focus more on current-year conditions. Specifically, their assessments deviated farther from the prior year assessment, and they documented and recalled more evidence that was inconsistent with the prior year assessment than did the preparers who were reviewed via e-mail. Additionally, it appears that the method of review influences preparer perceptions of accountability, with face-to-face review preparers
feeling more accountable to their managers than those anticipating an electronic review. Interestingly, in comparison to the no-review group, face-to-face preparers were significantly different with respect to their task performance and perceptions of accountability, while the electronic review group did not differ significantly. These findings suggest that it may not be prudent to view these two commonly utilized review formats (face-to-face and electronic review) as wholly interchangeable or equivalent substitutes for one another.

The remainder of this paper is organized as follows. The next section discusses the background and related research and develops the hypotheses. Sections III and IV present the method and results, respectively. Section V offers conclusions and implications.

II. BACKGROUND AND HYPOTHESIS DEVELOPMENT

During financial statement audits, the workpapers describing the work performed, methods used, and conclusions drawn by a preparer are subject to review by a supervising auditor (Emby and Gibbins 1988; Agoglia et al. 2002). Among other purposes, workpaper review serves to ensure the adequateness of procedures performed and appropriateness of conclusions drawn (AICPA 1978). Though a critical element of an audit, the significant allocation of audit resources to the review process has prompted firms to carefully examine workpaper review in an effort to streamline their audit processes (Rich et al. 1997). In addition, due to recent technological enhancements, firms have at their disposal alternative methods of review, such as electronic review of online workpapers.

While alternative review methods are employed in practice, prior research has typically focused on examining the effects of the expectation of workpaper review versus no review (e.g., Lord 1992; Kennedy 1993; Hoffman and Patton 1997; and Tan and Kao 1999). For example, Johnson and Kaplan (1991) found that auditors expecting a review of their inventory
obsolescence assessments exhibited greater consensus and self-insight than those who were told their responses would remain anonymous, while Koonce et al. (1995) found that auditors expecting to have their audit planning decisions reviewed recorded more information in their planning memoranda than those who expected their decisions would remain anonymous. Thus, prior research indicates that the anticipation of a review can affect preparer judgments when compared to a no review condition. However, in practice, auditors who prepare workpapers expect that they will be reviewed (Johnson and Kaplan 1991; Ismail and Trotman 1995). Moreover, these auditors are typically aware of how that review will be conducted (see footnote 2). Thus, since prior studies were designed to examine the effects of an expectation of review vs. no review, it is difficult to ascertain if firms might be able to utilize different forms of review as tools to elicit changes in preparer performance. The current study extends the literature by examining the effects that alternative methods of review may have on preparers’ performance.

**Alternative Methods of Review**

Based on a survey of practicing auditors, two alternative methods of review were selected for investigation: face-to-face review and electronic review. Face-to-face review occurs when the preparer meets with the reviewer in person. During this meeting, the reviewer relays his or her review notes to the preparer and the preparer initially responds to these concerns. In practice, this more traditional approach is generally interactive and discussion-filled, and often requires preparers to respond “on-the-spot” to reviewer inquiries, which may induce added stress on the preparer. Conversely, electronic review involves preparers receiving and responding to review notes via e-mail. This approach allows the preparer time to craft an initial response to the reviewer and potentially manage reviewer perceptions of his or her performance. In practice, use
of electronic review has risen with the proliferation of online workpapers. An advantage of this method is that it permits reviewers to review several jobs concurrently at their offices (where the majority of the firm’s resources are located) or from remote locations, reducing the time spent traveling between clients and the necessity to coordinate schedules with preparers (Shumate and Brooks 2001). However, the often rich and detailed interactions between reviewer and preparer may be lost. These differences between face-to-face and electronic communication have the potential to affect judgment performance (e.g., Baltes et al. 2002; Kachelmeier and Towry 2002).

The Heuristic-Systematic Model (HSM) suggests that contextual features of a judgment affect how an individual processes information (Chen and Chaiken 1999). HSM predicts two modes of information processing by which individuals make judgments (Chaiken 1980; Chaiken 1987; Chaiken et al. 1989; Chen and Chaiken 1999). Systematic processing involves the analytical and comprehensive treatment of judgment-relevant information, while heuristic processing involves the activation of judgmental rules (or “heuristics”) which help to process cues more easily (Chen and Chaiken 1999). Thus, relative to systematic processing, heuristic processing requires less cognitive effort. The choice between the two modes of processing depends on contextual features of the judgment. Decision makers attempt to balance the desire to minimize cognitive effort while maintaining a required level of confidence in their decision (Beach and Mitchell 1978; Payne et al. 1993; Chen and Chaiken 1999). HSM predicts that systematic processing is more likely when accuracy and confidence in judgment are the overriding concerns, while heuristic processing is more likely when time constraints and conservation of effort are the overriding concerns (Chen and Chaiken 1999). Thus, more systematic processing is expected when an individual places an emphasis on judgment.
effectiveness, while an emphasis on judgment efficiency should result in more heuristic processing.

Consistent with HSM, prior accountability research indicates that, under certain circumstances, higher levels of accountability can lead individuals to place greater emphasis on accuracy, use more systematic modes of processing information, and increase time and cognitive effort expended (e.g., McCallister, Mitchell, and Beach 1979; Chaiken 1980; Tetlock 1983; Tetlock and Kim 1987; Ashton 1992; Kennedy 1993; Koonce et al. 1995). The differences between face-to-face and electronic review have the potential to affect preparer accountability perceptions (Kreitner and Kinicki 2001). When a preparer receives a review note, he or she has encountered a situation which may threaten to undermine the desired image the individual wishes to present to a reviewer (Schlencker 1980). Depending on the perceived severity of the situation, the individual may feel his or her image is more or less threatened (Schlencker 1980; Tetlock 1985). It is possible that face-to-face reviews (which require immediate, and in-person, oral responses) and electronic reviews (which allow the preparer time to assemble a response and transmit it in writing electronically) may differentially affect preparers’ perceptions of the severity of the situation. In turn, their perceived accountability for workpaper preparation may differ.

Peecher and Kleinmuntz (1991) suggest that preparers will anticipate accountability demands early in the judgment process and these anticipated demands are likely to influence how they plan and carry out their work. Given that the formal evaluation process of preparers rewards both effective and efficient workpaper preparation, it is typically necessary for preparers to attempt to balance the often conflicting goals of audit effectiveness and efficiency (Bierstaker and Wright 2001). It is likely that the form of the impending review will influence the relative weight placed on these two concerns. In order to minimize the possibility of negative
consequences of their review (e.g., image deterioration), face-to-face preparers may perceive that their workpaper should be of higher quality, and that they must have a better working knowledge of the evidence, prior to review. This is likely to result in higher quality pre-review conclusions and supporting documentation in the workpapers. For these preparers, the balance between audit effectiveness and efficiency may be tipped in favor of effectiveness (i.e., a more systematic processing approach) as audit budgets may be somewhat sacrificed in an attempt to arrive at the most appropriate conclusions and prepare a set of high quality workpapers. More effective workpapers may help minimize potential reviewer concerns in advance and help to ready the preparer for the impending face-to-face meeting with the reviewer. In contrast, preparers anticipating electronic review will have more time to develop their responses to reviewer comments. Thus, these preparers may feel that they can affect the evaluation of their work on the audit more by meeting their budgets than by expending further effort on their initial workpaper submissions (i.e., a more heuristic processing approach). For these preparers, the balance between audit effectiveness and efficiency may be tipped in favor of efficiency, as pre-review workpaper/judgment quality may be somewhat sacrificed in order to maintain a desired audit budget. Therefore, the following hypotheses are tested:

**H1:** Preparers anticipating a face-to-face review will consume more of their time budgets than will preparers anticipating an electronic review.

**H2a:** Preparers anticipating a face-to-face review will prepare more effective workpapers than will preparers anticipating an electronic review.

**H2b:** The assessments of preparers anticipating a face-to-face review will deviate less from expert assessments than those of preparers anticipating an electronic review.

One of the most prevalent heuristics in public accounting is “anchoring” on prior year workpapers (Libby 1981; Wright 1988). Commonly referred to in practice as SALY (or Same As
Last Year), prior research has demonstrated the use of the anchoring heuristic in auditing contexts (e.g., Joyce and Biddle 1981; Butler 1986). Using last year’s workpapers and audit plans as the basis for the current year provides obvious opportunities for efficiency (e.g., using the same procedures as last year to test an inventory reserve). However, when conditions at the client have changed from the prior year, a SALY approach, while efficient, will not appropriately reflect those changes (Joyce and Biddle 1981; Tan 1995).

Prior research suggests that the expectation of review may result in increased attention to evidence that is inconsistent with prior year conclusions (Tan 1995). Using a financial viability task, Tan (1995) found that review awareness led to (a) better recall of current year evidence that was inconsistent with prior year data and (b) assessments that deviated farther from prior year assessments. Although that study did not investigate the effects of alternative review methods, it did show that the presence of a review can affect reliance on prior year workpapers.

The Heuristic-Systematic Model predicts a more systematic processing mode for face-to-face preparers and a more heuristic processing mode for electronic review preparers, given the characteristics of these judgment contexts. Further, HSM asserts that systematic processing leads to greater effort and greater attention to judgment-relevant information (i.e., current year evidence) than heuristic processing (Chen and Chaiken 1999). Thus, when greater emphasis is placed on effectiveness as a result of the anticipated method of review (i.e., there is a greater cost of committing a judgmental error), preparers may tend to focus more on current year evidence (as opposed to the prior year workpapers) than when the emphasis is on efficiency. In contrast, when efficiency is considered of increased importance relative to pre-review workpaper effectiveness, preparers may be more likely to rely on prior year workpapers and conclusions in an effort to conserve their budgets. It is therefore posited that, under changing client conditions, the current year assessments of preparers anticipating a face-to-face review will deviate farther
from prior year assessments than those anticipating an electronic review. The increased attention to more relevant evidence suggested by HSM should result in greater documentation and recall of that evidence. Additionally, given (a) their potential deviation from prior year conclusions due to changes in client conditions and (b) that they expect to be questioned in person by their reviewers, preparers anticipating a face-to-face review may perceive a greater need to have evidence which is inconsistent with prior year conclusions (but supports their current year assessments) readily available in their workpapers and retrievable from memory. This would be less necessary if the preparer were expecting an electronic review in which there would be time to reexamine any items the reviewer may have questioned. Thus, based on the discussion above, the following hypotheses are tested:

**H3:** The assessments of preparers anticipating a face-to-face review will deviate farther from prior year assessments than those of preparers anticipating an electronic review.

**H4a:** Preparers anticipating a face-to-face review will document relatively more evidence items that are inconsistent than consistent with prior year conclusions compared to preparers anticipating an electronic review.

**H4b:** Preparers anticipating a face-to-face review will recall relatively more evidence items that are inconsistent than consistent with prior year conclusions compared to preparers anticipating an electronic review.

III. METHOD

Forty-five practicing auditors from a large international public accounting firm participated in this study. Prior research and discussions with audit managers revealed that audit seniors would be familiar with the various stages of assessing the financial viability of their clients (e.g., Libby and Trotman 1993; Rau and Moser 1999). Auditors participating in this study were audit seniors with about 3 years experience on average.  

10
**Experimental Task**

Participating auditors completed a preliminary going concern evaluation task within a computer-based instrument. Participants received a diskette, along with a set of instructions for running the simulation program. Within the simulated audit exercise, participants were provided with detailed instructions pertaining to the task, relevant authoritative guidance, prior year workpapers, current year audit facts, and current year workpapers (which required them to make and document their preliminary conclusions).

The going concern task and evidence presented to participants were adapted from Tan (1995) and Kida (1984). Prior year workpapers documented a conclusion regarding the client’s going concern assumption, along with a corresponding memo summarizing the important evidence. The memo, presented in paragraph form, contained 10 items which were largely positive with respect to the financial condition of the client (see Tan 1995). The prior year preparer’s preliminary conclusion regarding the reasonableness of the client’s going concern assumption was presented in the prior year workpapers on a 15-point scale (-7 to +7, with endpoints labeled “definitely not reasonable” and “definitely reasonable,” respectively). Consistent with Tan (1995), this conclusion was favorable, with the prior year preparer having indicated that the “going concern assumption appears reasonable” and assessed the reasonableness of the assumption at +4 (labeled “fairly likely to be reasonable”). Current year audit facts included 10 items that supported the going concern assumption, 10 items that undermined the going concern assumption, and 10 irrelevant items. This current year evidence reflected a decline in client financial conditions (see Tan 1995). Similar to Cohen et al. (2000), auditor performance under declining financial conditions was investigated due to the increased risk associated with audit decisions under such conditions (e.g., the failure to modify the audit opinion prior to client bankruptcy).
Participants were asked to prepare a current year workpaper providing a preliminary audit conclusion (along with supporting documentation) regarding the reasonableness of the going concern assumption of the hypothetical client. Preparers were randomly assigned to one of three conditions: face-to-face review, electronic review, or no-review. Those in the two review conditions were told their work would be reviewed by an audit manager who would (a) have access to the prior year workpaper along with their current year workpaper and (b) evaluate them based on their combined audit effectiveness and efficiency after all review notes were addressed and their percentage of budget utilized reported. Preparers in the face-to-face review group were informed that they would meet in person with their reviewer to initially discuss any review notes regarding their going concern evaluation workpaper. Electronic review preparers were informed that all correspondence with their reviewers (including receipt of review notes and the preparer’s response to their reviewer) would take place over e-mail. An on-screen prompt required preparers in the two review conditions to acknowledge their review condition prior to beginning their going concern evaluation. Preparers in the no-review group were informed that their responses were confidential. A manipulation check revealed that preparers understood and anticipated their respective review conditions (i.e., face-to-face review, electronic review, or no review).11

Prior audit research investigating accountability has typically compared responses of subjects anticipating a review with those expecting their work to remain anonymous (e.g., Kennedy 1993; Glover 1997; Tan and Kao 1999). The no-review group serves as a baseline from which to compare the two review groups and allows us to gain insight into how these two methods of review that are currently used in practice (face-to-face and electronic review) are perceived by auditors, relative to an experimental condition present in many accountability studies.
Preparers were informed of their budget for the preparation of the workpapers (20 minutes). The budget began to count down after preparers read the task instructions. The “percentage of budget remaining” was displayed in the top right corner of each screen. If a preparer went over budget, a negative percentage was displayed on the “percentage of budget remaining” clock. All preparers had access to relevant excerpts from Statement on Auditing Standards No. 59 (AICPA 1988) and viewed identical prior year going concern evaluation workpapers and current year audit facts. Preparers were able to access this data while preparing their workpapers. Current year audit facts were randomized to control for order effects.

After viewing the prior year workpapers and current year audit facts, preparers provided their assessment of the reasonableness of the going concern assumption for the current year on a scale identical to the one on which the prior year conclusion was presented to them. They then provided documentation to support their conclusion and preparers in the two review conditions signed-off on their workpapers. This ended the timed portion of the task. Preparers then answered a series of confidential case-related questions, including a self-assessment of perceived accountability. After a distraction task, preparers were asked to recall as many of the current year audit facts as they could, and then answered some demographic questions. Preparers in the two review conditions later met or corresponded via e-mail with a reviewer.

IV. RESULTS

For clarity of presentation, the discussion of results is structured to focus on the hypotheses, which relate only to the face-to-face and e-mail review conditions. Comparisons of the review groups to the no-review condition are presented in the Additional Analyses section.
Efficiency and Effectiveness

Efficiency was measured in terms of the percentage of budget (20 minutes) consumed. Table 1 reveals that, consistent with H1, preparers in the face-to-face group consumed 141.67% of their budgets, significantly more than preparers in the e-mail review condition (101.00%, p = .012). More specifically, preparers in the face-to-face group spent more time in the actual workpaper preparation stage than did preparers facing an e-mail review (14.27 minutes vs. 8.53 minutes, p < .005). Presumably, this extra time (representing over 28.7% of the budget) was spent in an effort to prepare a higher quality workpaper. In addition to taking more time to perform their task, face-to-face preparers indicated in a post-experimental questionnaire that they were more likely to request additional budgeted time from their manager. On an eleven-point scale from 0 (“not at all likely [to request more time]”) to 10 (“extremely likely [to request more time]”), the self-reported mean for the face-to-face review group was significantly greater than that of the e-mail review group (means = 7.80 and 6.07, respectively, p = .032).

Three experts from different offices of the same accounting firm evaluated the effectiveness of the workpapers prepared by the participating auditors. These experts had an average of 10.55 years of audit experience. On an eleven-point scale from 0 (“strongly disagree”) to 10 (“strongly agree”), the experts expressed their agreement with a statement indicating that the preparer’s workpaper was effective. The average rating was used to create an overall workpaper effectiveness rating for each preparer. Hypothesis 2a predicted that the face-to-face group would prepare more effective workpapers than the e-mail review group. While the expert workpaper effectiveness ratings were in the expected direction, with experts rating the face-to-face group’s workpapers as more effective than e-mail (5.18 vs. 4.40), these differences were significant only at the p = .125 level (see Table 1). Participants’ self-reported measures
regarding effectiveness also were examined. These results indicated that the face-to-face group was more concerned with preparing an effective workpaper than was the e-mail review group. On an eleven-point scale from 0 ("not at all concerned with effectiveness") to 10 ("very concerned with effectiveness"), participants in the face-to-face group indicated higher levels of concern than the e-mail review group (means of 7.27 and 6.27 respectively, p = .087).

Similar to H2a, H2b predicted higher judgment quality for face-to-face preparers than e-mail preparers. Consistent with Tan (1995), judgment quality was measured by computing the absolute deviations of subjects’ judgments from the mean of an expert group. In our study, the face-to-face group’s mean going concern assessment (0.60) was nearly identical to that of the three experts (mean = 0.67), while the e-mail group mean (2.57) differed considerably from the experts’ assessment. Mean absolute deviations from the expert group were significantly smaller for the face-to-face group (2.60) than the e-mail group (3.51, p = .030, Table 1), providing support for H2b.

Deviation from Prior Year Workpapers

There were differences in preparers’ current year going concern assessments across the two review groups. Consistent with H3, preparers anticipating a face-to-face review deviated farther from the prior year assessment of +4 than preparers who would be reviewed via e-mail (mean deviations = 3.40 and 1.47, respectively, p = .013, Table 1). As stated previously, this deviation by the face-to-face group appears to be appropriate as their current year assessments were more in line with those of the audit experts.

Given that the prior year’s conclusion indicated that the “going concern assumption appears reasonable,” items that undermine the going concern assumption (i.e., negative items) are inconsistent with the prior year’s conclusion, while items that support the assumption (i.e.,
positive items) are consistent. Hypotheses 4a and 4b state that face-to-face preparers will document and recall relatively more evidence items that are inconsistent (i.e., negative) rather than consistent (i.e., positive) with the prior year workpaper’s conclusion compared to preparers anticipating e-mail review. The results support these hypotheses. On average, participants in the face-to-face group documented 2.87 more negative items than positive items, while those in the e-mail review group documented 0.40 more positive items than negative items (p = .007, Table 1). Also, face-to-face preparers recalled 2.40 more negative than positive items, while e-mail review preparers recalled, on average, 0.27 more positive than negative items (p = .009, Table 1).

Implicit in the expectations expressed in Hypotheses 3 and 4 is the assumption that e-mail review should result in greater heuristic processing (specifically, use of the anchoring heuristic) relative to face-to-face review. There is some evidence to support this notion. Consistent with a SALY (or anchoring) approach, the e-mail group appeared to be much more focused on the prior year than the face-to-face group. For example, the e-mail group spent significantly more time on prior year evidence and conclusions (2.73 minutes) than did the face-to-face group (1.80 minutes, p = .020). This occurred even though they spent significantly less time in total to complete the task (20.20 vs. 28.33 minutes). Additionally, while examining the current year and prior year evidence, the e-mail group focused less on current year data than did the face-to-face group, with 62% of their total evidence examination time spent examining current year data vs. 81% for the face-to-face group (p = .006). Also consistent with a SALY approach, the e-mail group spent less time preparing the current year workpaper (8.53 minutes) than did the face-to-face group (14.27 minutes, p = .001).
Additional Analyses

Accountability and Other Preparer Perceptions

As discussed earlier, the differences between face-to-face and electronic review may also affect preparer accountability perceptions (Kreitner and Kinicki 2001). Preparers were asked to indicate how accountable they felt “to [their] manager when performing the current year going concern evaluation.” They recorded their response on an eleven-point scale, with endpoints labeled “not at all accountable” (coded as 0) and “extremely accountable” (coded as 10). Preparers in the face-to-face review group reported feeling more accountable (mean = 8.27), on average, than those in the e-mail review group (mean = 5.20, p < .001, Table 2). Similar results were obtained when preparers were asked to indicate their motivation to complete the task, the mental effort expended on the task, how demanding they anticipated their review process would be, and the pressure they felt to perform well to impress their reviewer (see Table 2). Thus, it appears that the two forms of review do not evoke the same feelings of accountability, pressure, etc. from preparers.

The Influence of Accountability within HSM

Based on the reported effect of the method of review on the perceptions of accountability, we next consider the influence of accountability within the predictions of HSM. HSM (Chen and Chaiken 1999) suggests that contextual features of a judgment affect an individual’s desire for accuracy and confidence in judgment (concern for effectiveness). As concern for effectiveness increases, so should the use of more systematic information processing which, in turn, could result in differences in judgment. Thus, we might expect that preparers’ concern for effectiveness mediates the impact of review method on their going concern assessments. Further, prior
research indicates that, under certain circumstances, accountability can increase the desire for accuracy and confidence in judgment (e.g., Johnson and Kaplan 1991; Kennedy 1993; Tan 1995). This relationship suggests a three-path (two-mediator) mediation model presented in Figure 1 (MacKinnon and Tein 2003), which demonstrates the manner in which accountability perceptions and concern for effectiveness influence the going concern assessment. First, the method of review (face-to-face vs. e-mail) affects the preparer’s perception of accountability (path b₁). Perceptions of accountability affect the extent to which preparers are concerned with the effectiveness of their judgment (path b₂). Finally, the preparer’s concern for effectiveness influences the judgment (path b₃). While paths b₁ through b₃ show the indirect (mediating) effect of the method of review on the judgment, path b₄ indicates the direct effect.

To test this mediating relationship, we follow a well-established method for determining multiple mediation effects (e.g., MacKinnon and Dwyer 1993; MacKinnon et al. 1995; MacKinnon and Tein 2003). This method requires running three regressions and using the coefficients and standard errors from these regressions to build a test statistic for the mediating model. The three regressions are: perceived accountability on method of review (b₁); concern for effectiveness on accountability (b₂); going concern assessment on both concern for effectiveness (b₃) and method of review (b₄). Paths b₁ through b₃ are all significant (all individual p-values < .05). The test statistic for the mediation model, a z-statistic, results in a p-value for the model of .035. In addition, it should be noted that the effect of the independent variable (method of review) is significant only at p = .052 once the mediators are taken into account (increased from p = .018 in the original, non-mediated model). This decrease in significance for the independent variable is indicative of a mediating effect (Baron and Kenny 1986), lending additional support
for the model presented in Figure 1. Thus, it appears that accountability and concern for effectiveness mediate the effect of review method on preparers’ going concern assessments.

**Comparison with No-Review Control Group**

To facilitate comparison with prior studies investigating the effect of the review process on auditor decision behavior, this study compares both experimental groups with a control group which did not anticipate a review. These comparisons resulted in a generally consistent finding

[Insert Table 3]

(see Table 3). Based on analyses of the dependent variables discussed in H1 through H4 as well as preparer perceptions of accountability, motivation, and effort, the face-to-face group was significantly different from the control group. However, the e-mail group was not significantly different from the control group. This may suggest that it is not merely an impending review that affects the decision processes of audit workpaper preparers, but rather the characteristics of face-to-face review (e.g., less initial response time, the physical presence of the other party, synchronization of communication, etc.). Additionally, it may be that, for this type of task, an e-mail review was unable to increase the motivation, effort put forth, feeling of accountability, etc., beyond that already experienced by auditors who were not anticipating a review. Prior studies have shown that auditors in non-reviewed conditions generally report a relatively high level of motivation and effort exerted, despite the fact that their responses would remain confidential (see, e.g., Kennedy 1993; Glover 1997; Tan and Kao 1999). Indeed, Anderson and Maletta (1999) suggest that auditors bring a certain level of accountability to audit tasks irrespective of accountability manipulations.21
V. DISCUSSION AND CONCLUDING REMARKS

Due to technological advancements such as online workpapers and e-mail, audit firms have at their disposal alternative methods of workpaper review that they did not have in the past. While, in practice, preparers typically know they will be reviewed and know the form their review will take, prior research has focused on comparing the judgments and judgment processes of auditors who were aware their work could be reviewed with auditors who expected their work would remain anonymous (e.g., Johnson and Kaplan 1991; Lord 1992; Kennedy 1993; Tan and Kao 1999). This study extends the literature by demonstrating that the form of review chosen by reviewers can influence preparer performance.

The results of this study are consistent with the Heuristic-Systematic Model which predicts, in this context, greater emphasis on effectiveness (relative to efficiency) for face-to-face preparers compared to electronic review preparers. Results indicate that, compared to an electronic form of review (which was conducted via e-mail), preparers anticipating a face-to-face review were more concerned with audit effectiveness, produced higher quality judgments, spent more time preparing their current year workpapers, were less efficient at their task, were less likely to be influenced by prior year workpapers, and felt more accountable. These findings suggest that these alternative review formats affect preparers’ preferences for the often conflicting goals of audit effectiveness and efficiency. Specifically, an impending face-to-face review appeared to cause preparers to utilize a more systematic approach and be more concerned with the pre-review effectiveness of their workpapers, while electronic review appeared to cause preparers to utilize a more heuristic approach and place more value on efficiency. As a result, it may not be prudent to view these two commonly utilized review formats as wholly interchangeable or equivalent substitutes for one another.
To facilitate comparisons to prior research and to contribute to the understanding of the accountability construct in the audit environment, a no-review group was included in the experiment. Consistent with prior literature (e.g., Johnson and Kaplan 1991; Kennedy 1993; Glover 1997), preparers in the face-to-face group were significantly different than those in the no-review group with respect to their performance and perceptions of accountability. However, the electronic group did not differ significantly from the no-review group. Thus, these findings suggest that, more than simply an awareness of an impending review, the method of the review can affect preparer performance. Consequently, electronic review (a popular method currently used in practice) does not appear to create the same effects as more traditional “accountability” manipulations found in prior studies which typically involve the expectation of meeting in-person with a superior to justify conclusions (see e.g., Johnson and Kaplan 1991; Kennedy 1993; Glover 1997). In addition, and consistent with HSM, this difference in perceived accountability influenced the concern for workpaper effectiveness which, in turn, influenced the going concern judgment.

As with all research, our study’s limitations should be considered when evaluating the findings. For example, while our results suggest that electronic review can lower the quality of the going-concern judgment relative to face-to-face review, our study does not address whether this lower quality is below an acceptable level for this judgment. Also, we investigated the initial workpaper submission phase. It is possible that the efficiency gains of electronic review may attenuate by the time of the final workpaper submission (i.e., after review notes have been cleared) due to preparers being less concerned with initial workpaper effectiveness and, thereby, preparers and reviewers potentially spending greater time clearing review notes. This issue was beyond the scope of our study. It is also possible that differences in judgment quality between the two review methods may diminish at the final workpaper phase. However, recent research
suggests that differences in pre-review judgment are likely to persist after review (Agoglia et al. 2003). Additionally, in this study we examined only the two most common forms of review. There are other formats under which a review could be conducted such as by telephone or voice mail, and still others may evolve such as “live” electronic exchanges via the Internet.

The findings of this study have implications for practice and future research. Because the form of review may affect the behavior of preparers, both researchers and audit firms should consider the ramifications of alternative review methods with respect to the potential impact on preparers’ performance of the task at hand. For example, in a high risk audit environment where audit effectiveness is of greatest concern, reviewers may consider using face-to-face review. In contrast, electronic review may offer benefits beyond potential efficiency gains (such as enhanced coordination of information between audit team members) for lower risk tasks and tasks in which effectiveness is not significantly reduced by using this format. Future research could explore other possible effects of these two, and other, methods of review found in practice. For example, future research comparing alternative methods of review could examine issues such as task complexity (and other task characteristics) and preparer expertise to determine what methods may be most appropriate under differing circumstances. In addition, other important outcomes of the review process (e.g., training and evaluation of preparers) could be examined with respect to review formats. Also, the effect of alternative review methods on the perceptions and performance of reviewers (e.g., the effectiveness and efficiency of their review procedures) represents a fruitful area of research. Such research will further our understanding of the potential advantages and disadvantages of the possible methods of review for audit workpapers.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Audit Efficiency (percentage of budget consumed)</td>
<td>Mean (SD) 141.67 (57.00)</td>
<td>101.00 (28.92)</td>
<td>2.464</td>
<td>0.012</td>
</tr>
<tr>
<td>Workpaper Effectiveness (expert assessment)</td>
<td>Mean (SD) 5.18 (2.04)</td>
<td>4.40 (2.16)</td>
<td>1.175</td>
<td>0.125</td>
</tr>
<tr>
<td>Judgment Quality (deviation from expert criterion)</td>
<td>Mean (SD) 2.60 (1.37)</td>
<td>3.51 (1.17)</td>
<td>-1.963</td>
<td>0.030</td>
</tr>
<tr>
<td>Deviation from Prior Year Assessment</td>
<td>Mean (SD) 3.40 (2.72)</td>
<td>1.47 (1.51)</td>
<td>-2.408</td>
<td>0.013</td>
</tr>
<tr>
<td>Relative Documentation (positive minus negative items)</td>
<td>Mean (SD) -2.87 (3.58)</td>
<td>0.40 (3.11)</td>
<td>-2.665</td>
<td>0.007</td>
</tr>
<tr>
<td>Relative Recall (positive minus negative items)</td>
<td>Mean (SD) -2.40 (2.69)</td>
<td>0.27 (3.06)</td>
<td>-2.534</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Audit Efficiency is measured as the percentage of budget (20 minutes) utilized. Workpaper Effectiveness was determined by averaging the responses of three experts to a statement indicating that the preparer’s workpaper was effective. Responses were coded from 0 to 10, with the lower and higher endpoints labeled “strongly disagree” and “strongly agree,” respectively. The absolute deviations of each preparer’s assessment from the experts’ mean assessment (Judgment Quality) and from the prior year assessment (Deviation from Prior Year Assessment) were computed. Relative Documentation (Relative Recall) was computed as the number of positive minus negative items documented (recalled).

Due to homogeneity of variance issues, the t tests conducted for Audit Efficiency and Deviation from Prior Year Assessment did not use pooled variances. All tests are one-tailed.
### TABLE 2
Accountability and Other Preparer Perceptions

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-perceptions of:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accountability to reviewer</td>
<td>Mean (SD)</td>
<td>8.27 (1.39)</td>
<td>5.20 (2.34)</td>
<td>4.372</td>
</tr>
<tr>
<td>Motivation to perform well</td>
<td>Mean (SD)</td>
<td>7.80 (1.78)</td>
<td>5.13 (1.92)</td>
<td>3.941</td>
</tr>
<tr>
<td>Mental Effort expended on task</td>
<td>Mean (SD)</td>
<td>8.13 (1.85)</td>
<td>5.20 (1.66)</td>
<td>4.580</td>
</tr>
<tr>
<td>How demanding the anticipated</td>
<td>Mean (SD)</td>
<td>6.80 (2.27)</td>
<td>5.13 (2.39)</td>
<td>1.958</td>
</tr>
<tr>
<td>review would be</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pressure to impress reviewer</td>
<td>Mean (SD)</td>
<td>7.33 (1.91)</td>
<td>4.33 (2.74)</td>
<td>3.473</td>
</tr>
</tbody>
</table>

^aResponses were coded from 0 to 10, with the lower and higher endpoints labeled “not at all accountable” and “extremely accountable,” respectively, for the accountability assessment. Endpoint label pairings for the other self-perception assessments were: “not at all motivated” and “extremely motivated;” “no mental effort” and “extreme mental effort;” “not at all demanding” and “extremely demanding;” and “no pressure” and “extreme pressure.”

^bDue to homogeneity of variance issues, the t tests conducted for Accountability and Pressure did not use pooled variances. All tests are one-tailed.
# TABLE 3
No-Review Control Group Comparisons

<table>
<thead>
<tr>
<th>Variablea</th>
<th>No-Review Control [n = 15]</th>
<th>Comparison to Face-to-Face Review</th>
<th>Comparison to E-Mail Review</th>
<th>ANOVA (all 3 conditions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audit Efficiency (percent of budget consumed)</td>
<td>Mean 98.67 (SD 28.94)</td>
<td>test stat 2.907</td>
<td>p value 0.017</td>
<td>0.158</td>
</tr>
<tr>
<td>Workpaper Effectiveness (expert assessment)</td>
<td>Mean 3.76 (SD 1.00)</td>
<td>test stat 2.150</td>
<td>p value 0.025</td>
<td>0.974</td>
</tr>
<tr>
<td>Judgment Quality (deviation from expert criterion)</td>
<td>Mean 3.54 (SD 1.54)</td>
<td>test stat -1.871</td>
<td>p value 0.068</td>
<td>-0.043</td>
</tr>
<tr>
<td>Deviation from Prior Year Assessment</td>
<td>Mean 1.80 (SD 2.37)</td>
<td>test stat 2.043</td>
<td>p value 0.047</td>
<td>-0.405</td>
</tr>
<tr>
<td>Relative Documentation (positive minus negative items)</td>
<td>Mean 0.67 (SD 4.78)</td>
<td>test stat -2.489</td>
<td>p value 0.017</td>
<td>-0.188</td>
</tr>
<tr>
<td>Relative Recall (positive minus negative items)</td>
<td>Mean 0.00 (SD 3.78)</td>
<td>test stat -2.048</td>
<td>p value 0.047</td>
<td>0.228</td>
</tr>
<tr>
<td>Accountability</td>
<td>Mean 4.67 (SD 2.19)</td>
<td>test stat 4.891</td>
<td>p value &lt;0.001</td>
<td>0.725</td>
</tr>
<tr>
<td>Motivation to perform well</td>
<td>Mean 4.67 (SD 2.44)</td>
<td>test stat 4.151</td>
<td>p value &lt;0.001</td>
<td>0.618</td>
</tr>
<tr>
<td>Mental Effort expended on task</td>
<td>Mean 4.40 (SD 1.81)</td>
<td>test stat 5.773</td>
<td>p value &lt;0.001</td>
<td>1.237</td>
</tr>
</tbody>
</table>

a Audit Efficiency is measured as the percentage of budget (20 minutes) utilized. Workpaper Effectiveness was determined by averaging the responses of three experts to a statement indicating that the preparer’s workpaper was effective. Responses were coded from 0 to 10, with the lower and higher endpoints labeled “strongly disagree” and “strongly agree,” respectively. The absolute deviations of each preparer’s assessment from the experts’ mean assessment (Judgment Quality) and from the prior year assessment (Deviation from Prior Year Assessment) were computed. Relative Documentation (Relative Recall) was computed as the number of positive minus negative items documented (recalled). Preparers in the no-review group were asked how accountable they felt when performing the task. Given that these preparers understood that their workpapers were confidential, no reference to a reviewer was made in the
prompt for the Accountability assessment. Responses were coded from 0 to 10, with the lower and higher endpoints labeled “not at all accountable” and “extremely accountable,” respectively. On similar scales, these preparers were asked how Motivated they were to perform well and how much Mental Effort they expended on the task, with endpoints labeled “not at all motivated” and “extremely motivated” and “no mental effort” and “extreme mental effort,” respectively. F statistics are reported for the overall ANOVAs (which include all three preparer conditions), and t statistics for comparisons between two groups. Due to homogeneity of variance issues regarding comparisons between the face-to-face and no-review groups, the t tests conducted for Audit Efficiency, Workpaper Effectiveness, and Accountability did not use pooled variances. With respect to e-mail and no-review group comparisons, the t test conducted for Workpaper Effectiveness did not use pooled variances. All tests are two-tailed.
FIGURE 1
Multiple Mediation Model:
Perceived Accountability and Concern for Effectiveness as Mediators

$b_1$ through $b_4$ are the estimated coefficients from the following regression equations:

Accountability Perception = $\beta_{01} + \beta_1$(Method of Review) + $\epsilon_1$
Concern for Effectiveness = $\beta_{02} + \beta_2$(Accountability Perception) + $\epsilon_2$
Going Concern Assessment = $\beta_{03} + \beta_3$(Concern for Effectiveness) + $\beta_4$(Method of Review) + $\epsilon_3$

To determine the significance of the mediating relationship (paths 1 through 3), a z-score is calculated using the coefficient values and their standard errors. We calculate the variance of the mediating effect ($S_{b_1b_2b_3}$) using a first order Taylor series estimate (multivariate delta method) (Bishop et al. 1975; MacKinnon and Tein 2003):

$z = (b_1, b_2, b_3)/S_{b_1b_2b_3} = 1.813$ (p = .035)
REFERENCES


FOOTNOTES

1. It should be noted that a major purpose of these prior studies (e.g., Johnson and Kaplan 1991; Lord 1992; Kennedy 1993; and Tan and Kao 1999) was to investigate the effects of accountability on audit judgment. Therefore, low accountability conditions in these studies were typically operationalized to emulate conditions present in earlier studies (i.e., non-accountable), and not necessarily to replicate conditions found in practice.

2. In a post-experimental questionnaire, participants indicated their agreement with the statement that, in practice, they are generally aware of the method their review will take (mean response = 5.00 on a 1 to 6 scale, with the lower and higher endpoints labeled “strongly disagree” and “strongly agree,” respectively).

3. The going concern assumption, made for the purposes of financial reporting, refers to the expectation that an entity will continue operations (i.e., will not fail) in the absence of significant information to the contrary (AICPA 1988). Statement on Auditing Standards (SAS) No. 59 (AICPA 1988) states that, ordinarily, “information that significantly contradicts the going concern assumption relates to the entity’s inability to continue to meet its obligations as they become due without substantial disposition of assets outside the ordinary course of business, restructuring of debt, externally forced revisions of its operations, or similar actions . . . The auditor has a responsibility to evaluate whether there is substantial doubt about the entity’s ability to continue as a going concern for a reasonable period of time, not to exceed one year beyond the date of the financial statements being audited” (AICPA 1988, 1).
4. Ismail and Trotman (1995) manipulated discussion in the review process and found that discussion increased the number of plausible hypotheses generated by review groups on an analytical procedures task. However, no experimental manipulations occurred at the preparer level. All preparer subjects were informed that their judgments would be reviewed by another member of the firm. Their responses were used in both the discussion and no discussion review conditions. Thus, the effect of discussion during the review process on preparer judgments could not be determined.

5. Out of 14 audit managers surveyed, all 14 indicated that they utilized both of these methods of review. The managers’ mean estimates of the percentages of their total review time spent performing face-to-face reviews (“via discussion/in person”) and electronic reviews (“via e-mail/remotely”) were 34.6% and 62.5%, respectively. Additional forms of review mentioned by these managers included communications to preparers via telephone and voice-mail. Half of the managers indicated that they utilized face-to-face and electronic review exclusively. Six of the 14 managers were asked, informally, about methods they had used to review going concern evaluation workpapers. All six indicated that they had used both electronic and face-to-face methods to review this type of workpaper.

6. Similar to our face-to-face review condition, accountability manipulations in prior studies have typically involved some form of face-to-face interaction (or the anticipation of such an interaction) with the individual(s) to whom the subject was accountable (see, e.g., Johnson and Kaplan 1991; Kennedy 1993; Glover 1997).
7. For example, Baltes et al. (2002) illustrate that group performance is affected by the mode of communication between group members. While our preparer judgments involving a hierarchical review process differ significantly from the group decision processes discussed in Baltes et al. (2002), their results suggest that, for example, face-to-face communication tends to increase judgment effectiveness. Similarly, Kachelmeier and Towry (2002) find that the mode of communication during negotiation can affect negotiator behavior and perceptions of fairness. Although their context and communication conditions differ from those of our study on important dimensions (e.g., their context does not involve the perceptions of accountability present in a review setting, and their computerized negotiation condition does not allow participants to convince or persuade their counterparts), which likely results in fundamentally different participant motivations, their results support the notion that mode of communication can affect behavior.

8. For the purposes of this study, effectiveness is defined as the appropriateness of both the conclusions drawn by the auditor (utilizing an expert criterion) and the overall effectiveness of the workpaper (as judged by a group of experts). Consistent with Salterio (1994), efficiency is defined as the time taken to perform the task.

9. A survey of 14 audit managers confirmed that they expect a significantly quicker initial response to their comments when conducting a face-to-face review than for an electronic review.

10. There were no significant differences in mean months of audit experience between the groups (means = 32.27, 35.53, and 38.60, F = .860, p = .431, for the face-to-face, e-mail,
and no-review auditors, respectively). Also, there were no differences (all p’s > .50) between groups on other demographic variables (e.g., perception of the percentage of firms that fail, interest in the task, client size).

11. Participants in all review conditions were asked, “Do you expect that your workpaper will be reviewed (i.e., while completing your workpaper, were you anticipating a review)?” Participants who responded “yes” to this question were asked, “how do you expect your review will be conducted (e.g., in-person with your reviewer, electronically over e-mail)?” All responses were appropriate given the participant’s review condition (i.e., no-review preparers answered “no” to the first question, while preparers in the face-to-face and e-mail groups responded “yes” to the first question and indicated they expected to be reviewed face-to-face or via e-mail, respectively) except for one participant in the face-to-face review condition who 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.

12. Pretests involving three individuals with an average of about 4 years audit experience indicated that the budget was demanding, yet achievable, with pretest participants taking approximately 20 minutes, on average, to complete the task. In additional pretests, eighteen auditing students took 22.4 minutes, on average, to complete the task.

13. Significance levels of contrast tests presented in this study are one-tailed when expectations were directional, and two-tailed when no directional expectations were formed a priori. Also, in instances where homogeneity of variance is an issue, we perform tests of differences between groups which do not require the assumption of equal
group variances. The resulting p values of these tests using non-pooled variances are reported in the tables where noted. Further, homogeneity of variance issues may be due to outliers for certain dependent variables. Thus, analyses which exclude potentially extreme observations were also run in these cases. Results of analyses excluding potential outliers do not change any of the conclusions drawn in this study.

14. Prior to their involvement, the experts were informed that they would (1) examine evidence in order to assess the reasonableness of a hypothetical client’s going concern assumption and (2) evaluate the workpapers of others auditing the same client. In order to perform these tasks, the experts were provided with the same evidence presented to preparers. The three experts were blind to review condition and to preparer identity.

15. Tan (1995, p. 131) presents this measure as a “useful effectiveness benchmark in public accounting.”

16. The mean going concern assessment of these three experts (0.67) was consistent with that obtained from four audit partners (0.50), provided with identical prior year and current year information, in Tan (1995).

17. Given that in-person, synchronous interaction does not typically allow much time for reference, face-to-face preparers were not given the ability to reference the underlying evidence while providing their initial responses to reviewers. The face-to-face group was informed that they would not have “immediate access to the relevant audit evidence when meeting with [their] reviewer to discuss review notes.” In contrast, the e-mail group was told they would be able “to access all the relevant audit evidence when developing [their] responses to review notes after receiving them via e-mail.” To better understand the
mechanism driving our results (the form of interaction with the reviewer or the ability to access evidence), we created a modified face-to-face review condition in which, similar to the e-mail condition, preparers were informed they would be able “to access all the relevant audit evidence” during their meetings with reviewers. The 10 auditors (mean experience = 36.6 months) who participated in the modified face-to-face condition were similar to those in the main experiment, as they were not significantly different (all p-values > .40) on any of the descriptive variables measured (e.g., audit experience, estimate of percentage of firms that fail). The going concern assessments (mean = 0.70), deviations from experts (mean = 2.30), deviations from prior year assessment (mean = 3.30), consumption of budget (mean = 138.50), and perceptions of accountability (mean = 8.20) of preparers in this modified face-to-face review condition were in the expected directions and significantly different from those of the e-mail preparers (all p’s < .05), yet not significantly different than those of the original face-to-face group (all p’s > .50). These results are supportive of the notion that it is the anticipated method of interaction with the reviewer, and not the accessibility of evidence, that drive our results.

18. Evidence items were classified as positive (supports the going concern assumption), negative (undermines the going concern assumption), or irrelevant with respect to the going concern assumption based on the analysis performed by Tan (1992), who tested each item to determine its impact on a firm’s ability to remain viable. To examine preparers’ documentation and recall of evidence, two independent individuals and a researcher, all with over 4 years of audit experience, coded the documented and recalled evidence. Items were scored using a lenient gist criterion (i.e., preparers expressed the general meaning of the original item), with duplicate items and intrusions excluded (see
Tan 1995). Coders were blind to the experimental condition. The small number of differences between coders were subsequently reconciled.

19. These prior studies report results based on the manipulation of accountability, whereas we consider a mediating relationship of perceived accountability which is a measured variable in our study.

20. The z-statistic is an accepted method for testing the statistical significance of the mediating path (MacKinnon and Dwyer 1993). MacKinnon and Dwyer (1993) demonstrate (and MacKinnon and Tein (2003) further develop) how to calculate the z-statistic. The numerator is the mediated effect, which is the product of all of the coefficients for the mediating paths ($b_1b_2b_3$). The denominator is the standard error for the effect $S_{b_1b_2b_3}$, which is a function of the three standard errors and the coefficients. In this case the resulting z-score is 1.813.

21. It should be noted that the results regarding the e-mail review and no review conditions are not necessarily inconsistent with the results of prior accounting studies showing differences between review/accountable conditions and no review/non-accountable conditions. E-mail review is an inherently different form of review than those typically used in prior studies (e.g., Johnson and Kaplan 1991; Lord 1992; Hoffman and Patton 1997). Those studies generally employed a threat of review similar to, or stronger than, our face-to-face review condition (e.g., the anticipation of meeting in-person with a superior(s) to justify conclusions). Thus, one might expect comparisons between our face-to-face review and no review conditions to be similar to comparisons between review and no review conditions of prior studies, and our results are consistent with that
expectation. Given that our e-mail review condition, compared to conditions employed in prior studies, likely is viewed as a less threatening form of review, one might expect perceptions of accountability and the resulting effects to be mitigated. This, taken in conjunction with the relatively high baseline levels of accountability generally found with no review/non-accountable conditions (including our no review participants), likely should reduce one’s expectations of differences with respect to comparisons between our e-mail and no review conditions.