AUDITING: A JOURNAL OF PRACTICE & THEORY
Vol. 24, No. 2
November 2005
pp. 85-110

A Field-Based Analysis of Audit Workpaper Review

Neil L. Fargher, Diane Mayorga, and Ken T. Trotman

SUMMARY: This paper applies the models of Gibbins and Trotman (2002) and Rich et al. (1997a) in a public sector environment. We consider the factors impacting extent of review, level of stylization, and reviewer styles. We also examine the impact of reviewer rank on the above. In a field-based study we found that variation in review hours was due to both individual reviewer factors (reviewer style) and contextual factors (company size, time pressure, attitude of the reviewer to detail). Overall these results support the general robustness of the Gibbins and Trotman (2002) model to a public sector environment involving very experienced seniors and managers. Following the Rich et al. (1997a) framework we also provide evidence on the extent and types of stylization in the review process. Seniors and managers believed that the focuses of different reviewers differed substantially and that they stylized both the presentation and content of the working papers to accommodate these differences. While the evidence on their stylization was strong, their awareness of stylization by preparers for them was weaker. We also found significant differences between senior and manager reviewers including their extent of review, expectations of the perceived emphasis of the review level above, relative time spent on opinion formation and documentation, and the extent of their responses to reviewers' preferences. Finally, we provided a range of additional insights into reviewer style.

INTRODUCTION

For a number of decades the audit review process has been considered a key quality control mechanism of the audit process (Bamber and Bylinski 1982; Solomon 1987; Rich et al. 1997a). Drawing on the extensive experimental research on the review process (see review by Rich et al. 1997b), Gibbins and Trotman (2002) developed an interpersonal model of the audit manager as a reviewer. The model and field-based analysis suggest that a manager's conduct of the review is affected by the manager's expectation about the client, the preparer, and the partner, and the manager's own approach and circumstances. Following Gibbins and Trotman's (2002) interpersonal model we carry out a field-based analysis to investigate the conduct of a file review by experienced public sector senior- and manager-level auditors.

Neil L. Fargher is a Professor at Macquarie University; Diane Mayorga is a Lecturer and Ken T. Trotman is a Professor, both at The University of New South Wales.

We gratefully acknowledge comments on the paper from Elizabeth Carson, Bryan Cloyd, Amna Khalifa, Mark Peecher, Roger Simnett, Rick Tubbs, and two anonymous referees. We also acknowledge the comments of the participants at the presentation at The University of New South Wales.

*Submitted: October 2003
Accepted: May 2005*

Our study extends the Gibbins and Trotman (2002) research in the following ways. First, previous research on the review process has exclusively examined private sector financial audits. However, there is a trend toward public sector audit organizations adopting private sector audit standards. For example, the International Congress of Supreme Audit Institutions (INTOSAI)¹ uses International Standards on Auditing (ISAs) together with practice notes to be prepared by INTOSAI's Auditing Standard Committee. Practice notes include the extent of the ISAs' applicability and the adjustments required by a public sector auditor. As discussed in the next section of the paper, there are some significant differences between public and private sector auditors that have the potential to impact the audit process, including the review process. Thus, research is needed on a range of audit processes and quality control mechanisms in the public sector, including the review process, as inputs to INTOSAI's practice notes. In addition to the above practical reason, the public sector environment provides an opportunity to further test the Gibbins and Trotman (2002) model. Prior research has suggested the need for a theory to be tested through replications of the findings in different settings where the theory is expected to apply (Yin 1984; Spicer 1992).² They note that replication is important when research uses a field study methodology and each study examines a relatively small sample relative to the entire population.

Second, while Gibbins and Trotman's (2002) model of audit file review is centered on the audit manager, we examine the review process for both seniors and managers. Previous research has examined experience as a factor to explain between-reviewer variance. Research has found differences in time to complete the task (Ismail and Trotman 1995) and differences in types of errors detected (Ramsay 1994; Bamber and Ramsay 1997; Harding and Trotman 1999; Tan and Trotman 2003). We extend Gibbins and Trotman (2002) by considering both manager and senior review and their impact on the extent of review (review hours) as well as the focus of their review (documentation versus opinion formation). Understanding these differences in review between managers and seniors is important for developing appropriate training and decision aids for different levels within the firm.

Third, Tan and Trotman (2003) outline two conditions that need to be met before reviewers can provide adaptive responses to preparers' stylization attempts. One of these conditions involves reviewers being sensitive to such attempts and allocating appropriate attention toward the detection and correction of such stylization attempts. In this paper we examine how sensitive reviewers, of different levels of expertise, are to stylization attempts and extend Gibbins and Trotman (2002) by analyzing in greater detail how both preparers and reviewers learn about stylization. Given that persuasion knowledge is acquired over time as a result of experiencing persuasion attempts (Rich et al. 1997a) and that reviewers learn a variety of coping behaviors, we investigate whether both—anticipation of and coping with persuasion behaviors—will vary across managers and seniors. Rich et al. (1997a) suggest that such an understanding of these aspects of preparer and reviewer behaviors is an antecedent to making effectiveness evaluations of the review process.

Fourth, while Gibbins and Trotman (2002) considered review style, we extend this concept by examining the factors that affect this style. We provide evidence on the factors that impact reviewers' style including their choice of face-to-face review or written review notes and the emphasis on documentation, opinion formulation, and coaching. In addition,

¹ INTOSAI is the professional organization of Supreme Audit Institutions (SAIs) in countries that belong to the United Nations or its specialist agencies. SAIs play a major role in auditing government accounts and operations and in promoting sound financial management and accountability in their governments.

² The importance of replication is not however confined to field study methodology. Bamber et al. (2000) observe within the context of empirical archival studies that, because of the limited publication of replications, accounting researchers have little choice but to generalize from a small number of published studies.

we supplement the experimental studies on the types of errors that reviewers focus on (e.g., Ramsay 1994; Harding and Trotman 1999) and how the focus varies by rank (managers versus seniors).

These contributions are addressed using a questionnaire eliciting retrospective reports on public sector auditors' workpaper reviews of two clients. Retrospective reports have previously been used to investigate a range of issues including disclosures, negotiations, and the review process (Gibbins and Newton 1994; Gibbins et al. 2001; Gibbins and Trotman 2002). A potential bias with this method is that managers may describe how they believe the review might have happened instead of how it actually did happen. Following the guidelines of Gibbins and his colleagues, our design mitigates against this by asking about a large range of factors that could affect these items. Our results, using public sector auditors, support virtually all of Gibbins and Trotman's (2002) hypotheses related to extent of review. Our results include findings that reviewers anticipate stylization by preparers; this stylization relates to both presentation and work done; file preparers learn about reviewer styles in a variety of methods and the choice of method is impacted by familiarity; reviewers believe that there was substantial differences in review styles; and reviewers anticipate different review styles, which impacts content and presentation of working papers. In addition, we find some differences between senior- and manager-level review. Finally, we gain new insights into reviewer style and the factors that affect it.

The remainder of this paper is structured as follows. We present our theoretical framework and hypotheses in the next section. This is followed by the development and administration of our field questionnaire. The final sections provide the results and a discussion of these results.

THEORETICAL FRAMEWORK AND HYPOTHESIS DEVELOPMENT

Theoretical Framework

Our theoretical framework, based on Gibbins and Trotman (2002) (hereafter Gibbins and Trotman), incorporates two fundamental properties of audit file review. First, the Gibbins and Trotman interpersonal framework incorporates file preparers, audit managers, and audit partners involving continuous interactions between the parties enabling the formation of interpersonal beliefs and expectations among the three parties; advice/direction to the preparer by the manager and partner during the audit; the manager's file review of the preparer's work; and the partner's file review involving interaction between the partner and the manager. Second, the interpersonal framework incorporates expectations that play an important role up and down the hierarchy with key expectations including the preparer's expectations of the manager's review, the manager's expectations of the preparers that are likely to condition their review, and the manager's expectations of the partner's concerns that are likely to impact their reviews.

Gibbins and Trotman treat the manager as central to the review process and outline the following five components of their model centered on the manager: (1) manager's expectations about the client (for example, knowledge about risk, problem areas); (2) manager's expectations about the preparer (for example, quality of work, stylization of the file); (3) manager's expectations about the partner (for example, partner's review approach and focus); (4) manager's own approach to the review (manager style) and circumstances (for example, time pressure); and (5) manager's conduct of the file review (for example, quantity of review notes, time taken). Their model proposes that the first four components impact the fifth component. We adopt a similar framework but extend it to center on two parties, managers and seniors. We examine their expectations of those directly above and below them in the hierarchical process.

The persuasion framework developed by Rich et al. (1997a) provides considerable insights into the above expectations. Recognizing that audit working papers are an important means by which preparers demonstrate the quality of their work, Rich et al. (1997a) describe the review process from a persuasion perspective through the content and format of working papers. Following Rich et al. (1997a) we propose that there will be stylization of the working papers by the preparer (of content and/or format) in order to persuade the manager of the quality of the preparer's work. In addition, based on previous experience with the preparer (or discussion with other managers), the manager may expect stylization by the preparer and may have varying skills at coping with his/her stylization. Finally, the manager is co-composer of the file being sent to the partner and so, based on knowledge of the partner's preferences, may stylize the file to persuade the partner and enhance the manager's reputation (Rich et al. 1997a). In the present study, the level of stylization and the factors impacting it, are examined in an environment where there is often a long association between reviewer and reviewee.

The conduct of financial statement audits in this public sector organization we examined has many similarities to the private sector. The approach to the audit and the training materials were developed by a Big 4 audit firm. Similar to the private sector, audit tasks are performed in hierarchically structured teams and the audit decision-making environment is characterized by sequential and iterative review processes. A typical audit team includes of an engagement controller,³ an engagement manager, and a senior and a staff auditor. Discussion with senior staff⁴ supported the importance of the review process as a quality control mechanism.

On the other hand, there are a number of significant differences between public and private sector. First, there are differences in accountability that Kearns (1994) argues are the core differences between the public and private sector auditors. Public sector audit offices differ from their private sector counterparts because of their lack of profit orientation. Rather, their focus is on rendering services to parliament and to the community. This may impact auditing staffs' perceptions of their roles and duties within the audit team as well as their perceptions of their professional reputation and promotion criteria. Second, the demand for accountability in public sector audits is further highlighted by the increased public scrutiny by the media and special interest groups. Some audits are quite sensitive politically and therefore attract heightened media and special interest group coverage (Sinason 2000). This additional public exposure can influence the auditor's risk assessments and therefore affect the extent of audit testing and documentation included in the audit work papers. Third, the public sector audit offices have a tradition of maintaining audit team continuity for long periods. That is, there is the tendency to assign individuals for a number of years to the same audit engagement and the same engagement manager or engagement controller. In the private sector, staff generally work for a large number of superiors.⁵ Consequently by incorporating task familiarity as a variable we extend the Gibbins and Trotman model. Previous experimental research has found familiarity with the preparer can influence review activity (Asare and McDaniel 1996), but the impact of familiarity on reviewer persuasion behaviors has not been addressed in the literature. Fourth, given the additional emphasis on compliance with legislation, public sector auditors place

³ An engagement controller is similar to an audit partner role in the private sector.

⁴ Prior to conducting the study we discussed differences between private and public sector auditors with a number of public sector auditors. Subsequently, we interviewed two very senior public sector auditors with over 20 years audit experience (including at least ten years in the private sector Big 4 environment).

⁵ From our interviews we ascertained that seniors tended to work mainly for two managers and managers mainly worked for two engagement controllers. However, there is a policy of rotation off jobs after three to five years.

more importance on investigating the nature as well as the dollar value of an audit issue. This wider view of materiality can impact risk assessments, planning, evidence gathering, and the review process. Fifth, another key difference is the statutory independence of the Auditor-General. This is provided for in legislation by Parliament of each Australian jurisdiction in a number of ways including defining the scope of the Auditor-General's mandate, the appointment and removal of the Auditor-General, and the performance of his or her responsibility (Institute of Chartered Accountants of Australia [ICAA] 2003). Senior public sector auditors suggested, in interviews, that this is the purest form of professional independence and this perception impacts not only the relationship between the audited government organization and government audit office, but also the perceptions of public sector auditors regarding their roles and duties within the audit team.

Development of Hypotheses

We divide our hypothesis development into three sections. First, we address in a public sector environment the question of what factors impact the extent of review work. Second, we address a range of hypotheses and research questions on stylization and persuasion including reviewers' expectations of the person below and above them in the hierarchical process. Third, we address a range of research questions related to reviewer style. In each of these three sections we consider differences between managers and seniors.

Extent of Review

The Gibbins and Trotman model examined how client size, client relative riskiness, preparer quality, reviewer time pressure, prior review of the file, and reviewer style affected review pages and hours. Our basic research question in this section is whether the Gibbins and Trotman model generalizes to the public sector audit environment. Our belief is that this will be the case, based on the similarities between private and public sector financial audits outlined above. However, some important differences discussed earlier such as accountability, political sensitivity, familiarity, qualitative materiality, and level of independence could potentially lead to differences in the conduct of the review process.

RQ1: Are the factors affecting the extent of review in a public sector audit similar to those found by Gibbins and Trotman (2002) in the private sector?

In addition, we extend Gibbins and Trotman by incorporating three other variables in their model: rank of the reviewer, level of detail required by the reviewer, and familiarity with the reviewer. Managers are central to the review process (Gibbins and Trotman 2002), and we would expect that they spend more time on review than seniors. We test this in H1a.⁶

H1a: Managers' conduct of their file review involves more hours than senior reviews.

Rich et al.'s (1997a) model of the review process proposes that there will be stylization of working papers by the preparer to persuade the reviewer of the quality of the preparers

⁶ We also expect some other differences between manager and senior reviews that will be outlined in subsequent hypotheses where we consider awareness of reviewee's stylization (H2b), expectations of the reviewer (H2e), responses to the concerns and preferences of reviewers (H2f), and time spent on opinion formation and documentation (H3).

work. Similarly the reviewer, as a co-composer of the file being sent to the next level of review, is likely to stylize the file to persuade the partner/engagement controller and enhance the reviewer's reputation. Consequently, we would expect that both seniors and managers would increase the extent of review when the reviewer requires more detail.

H1b: Managers and seniors both increase the extent of review when the next level reviewer is considered to be more detail oriented.

Previous research has shown that familiarity with other team members has implications for reviewer's performance (Asare and McDaniel 1996) and assessment of subordinate's work (Tan and Jamal 2001). Gibbins and Trotman's descriptive evidence shows that reviewers spend time coaching and training staff during the review process. We would thus expect the extent of review to decrease as familiarity increases since the reviewee⁷ would be more likely to be aware of his or her reviewer's concerns and preferences. We therefore test the following hypothesis.

H1c: Familiarity with the reviewee is negatively associated with the extent of review.

Stylization and Persuasion

Rich et al. (1997a) describe the review process from a persuasion perspective. They suggest that preparers stylize working papers and reviewers receive working papers that contain messages intended to persuade reviewers about the appropriateness of the work performed and conclusions reached. For reviewers to be effective, reviewers must anticipate the persuasive behaviors of other audit team members (Rich et al. 1997a; Tan and Trotman 2003). Rich et al. (1997a) also suggest that the role of managing the file for a partner results in the manager becoming co-composer of persuasive working paper messages. In our study we envisage both seniors (who have done one level of review) and managers becoming co-composers. Seniors anticipate the focus of the manager's review and are likely to stylize the working papers accordingly. Furthermore, managers then anticipate the focus of the engagement controller's review and are likely to stylize the review accordingly. Here we first consider expectations of reviewers (managers and seniors) about reviewees (H2a, H2b, RQ2.1–RQ2.3). Second we consider reviewee (manager and senior) expectations about reviewers (H2c, H2d, H2e, and RQ2.4).

We expect the Rich et al. (1997) theory on persuasion to apply to a public sector environment. Our public sector participants have substantial audit and review experience that should allow them to develop persuasion knowledge that will impact their expectations about the reviewee's work and their expectations about their reviewer's approach to review. Our public sector setting is one in which you would potentially expect a greater anticipation by reviewers of stylization because of the longer association between reviewers and preparers in public sector audits.

H2a: Reviewers anticipate file stylization by reviewees (Gibbins and Trotman 2002, H2b).

⁷ We use the term reviewee to describe the person preparing the working papers for the senior to review and a senior whose work is being reviewed by a manager.

Rich et al. (1997a) note that reviewers learn about persuasion through a number of ways including reviewing files, discussions with other reviewers and knowledge acquired while they were preparers. Rich et al. (1997a) note persuasion knowledge is acquired over time as a result of experienced persuasion attempts. This suggests that managers and seniors will have different persuasion knowledge and therefore likely to have different expectations about the preparer and the preparer's work. We would expect that the greater persuasion knowledge of managers would make them more aware of and more likely to recognize persuasion attempts. Furthermore, Tan and Trotman (2003) demonstrate that auditors' responses to stylization attempts depend on auditor rank.

H2b: Managers are more likely to be aware of reviewees' stylization than seniors.

Gibbins and Trotman analyzed the comments of managers on stylization to determine why stylization occurred. The responses included the preparer learning from previous experience, the preparer learned stylization on the job, or the preparer was advised of the reviewer style at the planning stage. Here we extend Gibbins and Trotman by directly asking participants how they inform file preparers about their preferences. Reviewers have choices about whether they inform preparers of their preferences and if so, when they inform them. Finally, there is variation in reviewers' beliefs of whether preparers stylize for them and we consider what factors impact that perception. We consider the following three research questions, of which the latter two have not previously been addressed in the literature.

RQ2.1: How do reviewers inform file preparers about their preferences? (See Gibbins and Trotman 2002, Additional Analysis, Table 3.)

RQ2.2: What factors determine whether reviewers inform preparers of their preferred style at the planning stage?

RQ2.3: What factors impact whether reviewers believe that reviewees stylize for them?

We now consider expectations by managers and seniors about their reviewers. Following Gibbins and Trotman we consider the following two hypotheses:

H2c: Managers and seniors both perceive that reviewers differ in their approach to their review of the file received from them (Gibbins and Trotman 2002, H3a).

H2d: Managers and seniors both perceive that they stylize their reviews in accordance with their reviewers' approaches (Gibbins and Trotman 2002, H3b).

It should be noted that while the above two hypotheses replicate Gibbins and Trotman in a public sector environment, the data collection method, as described in the research methods section, is different in that the present study uses structured questions rather than open-ended questions, thus eliminating the need for coding. These responses also provided us with substantial additional information (130 compared to 37 explanations in Gibbins and Trotman).

We expect differences between managers and seniors in their expectations of their reviewer. Previous research has found that during the review process, managers adopt conclusion error templates while seniors adopt documentation error templates (Ramsay 1994; Tan and Trotman 2003). Given these templates we would expect seniors' perceived emphasis of reviewers to be more related to documentation/detail, while managers' emphasis would be more related to conclusion/big picture and critical issues (that is, conclusion-type issues).

H2e: Senior (manager) expectations of the reviewer will be more related to documentation (conclusion) issues.

The different review templates for managers and seniors (Ramsay 1994) may potentially impact the preparation of the file for the next level of review. While Tan and Trotman (2003) demonstrate that auditors' responses to stylization attempts depend on auditor rank, no research has looked at the impact of rank on reviewers' responses to the concerns and preferences of subsequent reviewees by the level above. However, we would expect that the better developed schemas of the managers would result in managers being less affected by the concerns and preferences of the reviewer than the seniors. That is, managers would have developed a review style over a longer period and would be less likely to change that style.

H2f: Seniors are more affected than managers by the concerns and preferences of reviewers.

Reviewer Style

In this section we examine a range of research questions on reviewer style identified (but not answered) in previous literature. We consider different approaches to review, including face-to-face versus written review notes (Rich et al. 1997b), and the consistency of these approaches. Prior research also shows that the important review objectives include: checking documentation, assessing opinion formulation, and coaching/training of staff (Gibbins and Trotman 2002; Solomon 1987). In this study we asked both seniors and managers to report, for each of their two cases, the percentage of time taken on checking documentation, assessing opinion formulation, and training/coaching. We then investigated whether these percentages vary between cases, across rank, and what other factors affect the proportion of review for each of these activities.

RQ3.1: What approach to review is preferred by reviewers (e.g., face-to-face, written comments only, etc.) and does this differ between seniors and managers?

RQ3.2: Is the auditor's approach to review (face-to-face, etc.) consistent between cases?

RQ3.3: What proportion of review time is spent by managers and seniors on checking documentation, assessing opinion formation and coaching/training of staff?

RQ3.4: Is the auditor's proportion of review on the above factors consistent between cases?

RQ3.5: What factors affect the proportion of review on each of the above objectives?

We also test a specific hypothesis related to differences in time spent on assessing opinion formation and checking documentation. Previous studies (Ramsay 1994; Tan and Trotman 2003) have illustrated the different knowledge structures of seniors and managers using recognition tests, and that this impacts performance (Tan and Trotman 2003). We would also expect that knowledge structures would impact the amount of time spent on checking documentation and assessing opinion formation by managers and seniors. As seniors (managers) knowledge structures are built around documentation errors (conceptual errors) (see Ramsay 1994; Tan and Trotman 2003) we would expect the managers to spend more time than seniors on opinion formation (conceptual errors) and less on documentation errors.

H3: Managers spend more time than seniors on assessing opinion formation and less time than seniors on documentation.

DEVELOPMENT AND ADMINISTRATION OF THE FIELD QUESTIONNAIRE

Participants

The questionnaire was completed by 18 seniors and 18 managers (62 percent response rate) from an Australian public sector audit organization. Table 1 summarizes the descriptive statistics for the sample of 72 organizations.⁸ Seniors averaged 10.2 (9.2) years of auditing experience (public sector auditing experience). On average, seniors had been doing file reviews for 5.1 years. Managers averaged 18.2 (16.4) years of auditing experience (public sector auditing experience). On average, managers had been doing file reviews for 10.4 years. While the managers have significantly more experience than seniors, there is no evidence of a significant difference between managers and seniors in their responses to the questions detailed on Table 1.⁹

Questionnaire Content and Administration

To investigate review issues in their natural setting, an extensive questionnaire consisting of ten sections over 15 pages was developed.¹⁰ A contact person in the public sector auditing organization distributed 27 questionnaires to managers and 31 to seniors. Additionally, a consent letter was emailed to the participants by a very senior member of the organization authorizing their participation, and distributed to participants by an assistant auditor general. Participants were assured that all individual replies would be kept completely confidential. The questionnaire requested that the participants not identify themselves, their division, or their clients. Participants returned the envelopes directly to the researcher. Respondents were given a gift certificate for AUS\$75 (U.S.\$40). A covering letter from the authors asked for frank responses and provided assurance of anonymity.

⁸ The sample includes 23 government departments and 49 state-owned enterprises or statutory authorities including universities. While the government departments tended to take more review hours ($t = 2.57$, $p = 0.01$), there was no significant difference in the measure of inherent risk between the government departments and other types of clients ($t = 0.95$, $p = 0.35$). The results reported are not influenced by inclusion of an additional indicator variable for government department.

⁹ There is a large variation in the hours taken to complete the review and the size (total assets) of the audit clients. The size variable was Winsorized at the 5 percent and 95 percent levels. The distribution of review hours and client size are both skewed and so the log of these variables is used in the subsequent analysis to mitigate the influence of outliers on the distribution of the regression residuals. Diagnostics were reviewed for potentially influential observations but no studentized residuals are in excess of 2.5.

¹⁰ Although the questionnaire was long there was no indication of fatigue. There were no signs of the final pages being completed hastily, the last page of demographics was completed by all participants, and the order variable for case 1 and 2 was not highly correlated with any other variables.

TABLE 1
Descriptive Statistics

Variables ^a	Full Sample Means	Seniors Means	Managers Means	Prob. of Difference ^b
Audit Experience (years)	14.583	10.982	18.194	0.002**
File Review Experience (years)	7.750	5.074	10.426	0.001**
Log (<i>REVIEW HOURS</i>)	3.187	2.651	3.725	0.001**
Log (<i>SIZE</i>)	4.709	4.198	5.220	0.073
Client's relative riskiness (<i>RISKINESS</i>)	5.042	4.833	5.250	0.296
Relative quality of file preparer	5.819	5.861	5.778	0.789
Relative quality of preparer's opinion formulation	5.486	5.278	5.444	0.775
Relative quality of preparer's documentation	5.472	5.389	5.555	0.578
Average of three preparer quality ratings (<i>AVGQUAL</i>)	5.593	5.593	5.593	1.000
Relatively how busy was manager	6.708	6.639	6.778	0.697
Deadline pressure compared to normal	6.458	6.222	6.694	0.149
Average of two reviewer time pressure ratings (<i>AVGTIME</i>)	6.583	6.431	6.736	0.331
Engagement partner's attitude (<i>ATTITUDE</i>)	5.208	5.250	5.167	0.854
Preparer adjusts file indicator variable (<i>PREPSTYF</i>)	0.556	0.556	0.556	1.000
Familiarity with the preparer (<i>FAMILIARITY</i>)	4.131	4.263	4.000	0.776

** Significant at $p < 0.01$ (one-tailed).

^a Definitions of variables:

REVIEW HOURS = number of review hours spent on financial audit;

SIZE = estimated gross assets winsorized at the 5 percent and 95 percent levels;

RISKINESS = rating of inherent riskiness of client compared to other clients;

AVGQUAL = average of three preparer quality ratings;

AVGTIME = average of two reviewer time pressure ratings;

ATTITUDE = next level of reviewer's attitude to working paper details compared to others; higher is more pedantic;

PREPSTYF = file preparer adjusted file to suit reviewer's preference for financial audit (Yes/No); indicator variable takes the value 1 for yes and 0 for no; and

FAMILIARITY = number of audit engagements auditor has worked with file preparer.

^b Probability of difference using a t-test for the difference of means.

With the assistance of some senior public sector auditors, the questionnaire was adapted from the questionnaire used by Gibbins and Trotman to the particular public sector organization's review setting. We had participants provide information on two separate audits, and for both cases to think of a recently completed set of working papers they reviewed where the audit fee was at least \$10,000.¹¹ Participants choose the first case without direction and for the second case they were asked to choose a file that had significantly more

¹¹ Advice from an Assistant Auditor General indicated that a low fee was required to capture audit engagements reviewed at the less experienced senior level. Higher fee engagements tend only to be reviewed at the manager level. Approximately 80 percent of this organization's fee revenues are generated from only 20 percent of their clients.

or less review time than the first case. Review hours were then analyzed in a repeated-measures design.¹² Participants were asked a series of questions about the first set of working papers chosen including: size (gross assets and employee numbers), type of client, time spent of review (main dependant variable), and a number of measures of prior work experience with other participants in the review process. Participants were then asked to circle the appropriate point on a nine-point scale to provide the following seven ratings for the particular audit: the quality of the work expected from the file preparer, how busy the reviewer was, extent of deadline pressure, quality of file preparer's opinion formation, quality of file preparer's documentation, inherent risk, and reviewer's attitude to working paper details. This was followed by a number of questions seeking descriptive evidence on the main objectives of the review process including the percentage of time spent on checking documentation, assessing opinion formation, and training/coaching of staff. A question also asked for the percentage of the review done face-to-face before written comments, face-to-face after written comments, and written comments only.

For each set of working papers selected by the participants we asked whether the file preparers adjust the file to suit the reviewer's style preferences. We then asked whether the adjustment was to the presentation of the working papers, the type of audit work done, or both; and how they informed the file preparer about their preferences.

The questionnaire also asked questions that were not dependent on the two cases but provided additional descriptive support for the study. To obtain further insights on stylization, participants described how the next level of review affects what they do in their review. Also managers (seniors) were asked to select three engagement controllers (managers) and to list any particular concerns or preferences that these reviewers have that affect the way files are prepared by them. They then choose from the following list to indicate how the concerns or preferences noted above affected the way they prepared the files: affects working paper content; affects working paper presentation; affects both content and presentation; no effect on file presentation or content. Finally, the last section requested demographic data and provided space for any additional comments from the participants. Both seniors and managers received the same series of questions in the same order.

In addition to a number of minor wording changes, our research instrument had the following differences to Gibbins and Trotman, which are shown in parentheses: there was a separate questionnaire for seniors and managers (managers only); the audit fee needed to be at least \$10,000 (\$50,000); client size was measured by gross assets (annual revenue); requested review hours (review hours and pages of review notes); four questions to address familiarity with the preparer and higher level reviewer (nil); time spent on checking documentation, assessing opinion formation, and training/coaching of staff (checking documentation and assessing opinion formation); after the question "Did any of the preparers adjust the file to your reviewing style? Yes/No" we asked "If so, did the file preparer: adjust presentation of working papers; adjust type of audit work done; both of the above" (no categories given; participants were asked to "please explain"); added questions on "How did you inform the file preparer about your preferences?" with four choices listed (no equivalent questions); seniors/managers asked to choose three managers/controllers and describe their concerns/preferences (choose four partners); asked to choose how their concerns/preferences affect the way they prepare the files, working paper content, working

¹² Data for both the financial statement audit and the compliance audit were collected. Only the former were analyzed in this paper. Participants were reminded in the questionnaire that the financial statement audit refers to all parts of the audit except for the compliance audit.

paper presentation, both content/and presentation, or no effect on presentation or content (describe their concerns/preferences).

Questionnaire Coding and Analysis

Given the difficulties of coding some of the narrative data in previous studies (e.g., Gibbins and Trotman 2002), we reduced the number of open-ended questions and where possible used categorical questions. There were two questions that needed coding: (1) How does the engagement controller/manager affect what you do in your review? (see Table 6); (2) Choose three engagement controllers/managers and describe below how any particular concerns or preferences each has that affect the way files are prepared for them (see Table 5, Panel A). This data was coded by two of the coauthors and the small number of differences was reconciled. Following the procedures in earlier studies (Gibbins and Newton 1994; Gibbins and Trotman 2002), the above coding was done from photocopied excerpts to avoid possible influence by other responses in the questionnaire.

The first part of the questionnaire contained repeated measures for the two cases and was analyzed using both univariate and repeated measures multiple regressions. The regressions incorporate a set of independent variables, including the size of the client (log of total assets to reduce the impact of outliers on the analysis), and ratings of relative risk, preparer quality, deadline pressure, and the partner's attitude to the review (pedantic or not).¹³ In this repeated measures model, following Gibbins and Trotman, we include an indicator variable (0,1) for each respondent as each auditor provides data for two audits. Because individuals are either managers or seniors we estimate coefficients for each group and then compare the average coefficients for managers and seniors. We also include an indicator variable for the first of the two cases recalled to mitigate the potential impact of any bias arising from the order in which the cases were recalled. Demographic variables could not be included in the repeated measures regressions because those variables were measured only once and statistically all individual differences are incorporated into the repeated measures.

RESULTS

Extent of Review

We begin our analysis on the factors affecting the extent of review by first considering RQ1, comparing the factors affecting the extent of review in a public sector audit with the results of Gibbins and Trotman in the private sector. Table 2 provides a full correlation matrix of our variable and Table 3 provides estimates for a multiple regression model of review hours described above. The univariate analysis is inconclusive, as review hours are significantly positively related both to company size and to client risk. The repeated measures regression shows that in the presence of all the other variables, and taking out the effect of individual differences represented by the repeated measures, four variables were significant: client size (larger clients are associated with more review time), time pressure,¹⁴ reviewer's attitude, and repeated measures indicator variables. Reviewer's attitude was not included in the Gibbins and Trotman model and is addressed in H1b.

¹³ While we have most of the determinants for review hours identified by Gibbins and Trotman (2002) we do not have a variable for having reviewed this file before because the high stability of auditors in the public sector results in almost all files having been reviewed before.

¹⁴ Participants were asked to rate how busy they were compared to other times during the year and to rate the deadline pressure compared to normal pressure using a nine-point scale. Because these two measures are highly correlated (Spearman correlation 0.69) an average of the two time pressure ratings was taken and it had a mean of 6.58.

TABLE 2
Spearman Correlations between Variables

Variable	HRS	SIZE	RISK	Q1	Q2	Q3	AVG QUAL	T1	T2	AVG TIME	ATTITUDE	PREP STYLE	FAMIL MGR.
HRS	1.00												
SIZE	0.59**	1.00											
RISKINESS	0.28*	0.53**	1.00										
Quality1 (Q1)	-0.08	-0.12	0.04	1.00									
Quality2 (Q2)	-0.04	-0.06	-0.11	0.58**	1.00								
Quality3 (Q3)	-0.01	-0.03	-0.05	0.72**	0.66**	1.00							
AVGQUAL	-0.03	-0.09	-0.04	0.88**	0.83**	0.90**	1.00						
TimePressure1 (T1)	-0.02	0.01	-0.07	0.04	0.06	0.05	0.06	1.00					
TimePressure2 (T2)	0.25*	0.10	0.05	0.11	0.14	0.08	0.14	0.69**	1.00				
AVGTIME	0.11	0.05	-0.00	0.08	0.11	0.08	0.12	0.92**	0.90**	1.00			
ATTITUDE	0.09	0.01	0.07	-0.00	-0.11	-0.13	-0.07	-0.39**	-0.14	-0.30*	1.00		
PREPSTYF	0.09	0.07	0.01	-0.14	-0.14	-0.10	-0.17	-0.05	-0.18	-0.11	0.05	1.00	
FAMILIARITY	0.09	0.02	0.09	0.13	0.09	0.06	0.11	-0.13	-0.03	-0.09	0.18	0.04	1.00
MANAGER	0.52**	0.20	0.13	-0.06	-0.03	0.03	-0.00	0.05	0.18	0.13	-0.08	0.00	-0.11
ORDER	0.04	0.21	0.16	0.23	0.23*	0.16	0.23	-0.00	-0.06	-0.03	-0.16	0.11	0.10
													0.00

*, **, Significant at $p < 0.05$ and $p < 0.01$, respectively (two-tailed test).

TABLE 3
Regression on Log of Review Hours^a

Independent Variables^b	Prediction	Coefficient (t-statistic)
Client size	(+)	0.176 (4.74)**
Client's relative risk	(+)	-0.130 (2.19)
Average quality	(-)	0.020 (0.26)
Average time pressure	(+, -)	0.177 (2.40)*
Reviewer's attitude	(+)	0.111 (2.47)*
Preparer adjusted file	(-)	0.165 (0.67)
Familiarity	(-)	0.017 (0.70)
Order variable	(.)	-0.042 (0.31)
Repeated measures indicator variables—managers	(.)	1.469 (8.98)**
Repeated measures indicator variables—seniors	(.)	0.561 (3.53)**
Number of observations		72
Regression Overall R ²		0.993
Regression Adjusted R ²		0.982
F-Statistic		87.76**

*, ** Significant at $p < 0.05$ and $p < 0.01$, respectively, (one-tailed).

^a Regression of log of review hours on the factors expected to influence the extent of review. As each auditor provided two cases we include a dummy variable for each manager and each senior. The average of the coefficients for managers and the average of the coefficients for seniors are reported in the table. The prediction is that managers will have higher hours spent on review than seniors.

^b Definitions of variables:

Log of Review Hours = log of the number of review hours spent on financial audit;
 Client Size = log of the estimated gross assets winsorized at the 5 percent and 95 percent levels;
 Client's relative risk = rating of inherent riskiness of client compared to other clients;
 Average Quality = average of three preparer quality ratings each measured on a nine-point scale;
 Average Time Pressure = average of two reviewer time pressure ratings each measured on a nine-point scale;
 Reviewer's Attitude = next level of reviewer's attitude to working paper details compared to others; higher is more pedantic; measured on a nine-point scale;
 Preparer Adjusted File = file preparer adjusted file to suit reviewer's preference for financial audit (Yes/No); indicator variable takes the value 1 for yes and 0 for no;
 Familiarity = number of audit engagements auditor has worked with file preparer;
 Order = an indicator variable taking the value 1 for case one and 0 for case two; and
 Repeated measures indicator variables = dummy variables for each manager and each senior. Two variables are excluded to allow model estimation. The average of the coefficients for managers and the average of the coefficients for seniors are reported separately.