Indian Accounting Review, Vol. 11, No. 2, 2007, pp. 1-21

The Working Paper Review Process: A Theoretical Model and Simulations

Hai Lu
Assistant Professor, Rotman School of Management
University of Toronto, Toronto, Canada
Email: hai.lu@rotman.utoronto.ca

Rajendra Srivastava*
Ernst & Young Professor
The University of Kansas
Lawrence, KS 66045, USA
Email: rsrivastava@ku.edu

And

Theodore J. Mock
Professor of Accounting
University of Southern California
Professor of Auditing Research
University Maastricht
Email: tmock@marshall.usc.edu

December 2007

*Corresponding author

The Working Paper Review Process: A Theoretical Model and Simulations

Abstract

This study presents a theoretical model of the working paper review process within the financial statement audit. The model is based on the concepts of reliability engineering with auditors functioning as the main control components. Characteristics of the model are investigated using simulation and differential analysis. Our simulation results show that the performance of a combined review team is more effective than an individual review; the review performance of a team consisting of an audit senior and an audit manager is better than that of the team consisting of two seniors; and the performance of an encompassing review differs from the performance of an error specialized review in a sequential working paper review process. The simulation results also show that a team review may not be an optimal choice when the expected litigation cost is of the same order as the expected operating cost of each review. The fact that managers are expected to detect more conceptual errors and irregularities leads to lower expected litigation cost. Consequently, assurance firms may be better off by choosing a senior-manager review even if the overall expected reliability from a senior-senior review is identical to the review from a senior-manager review. The differential analysis indicates that the parameters representing the motivation and the effort level of two reviewers are not of the same importance in the working paper review process and that the increase in the quality of working paper input will not automatically increase the reliability of the output.

Key Words: Working Paper Review, Reliability of Information, Control, Simulation.

The Working Paper Review Process: A Theoretical Model and Simulations

I. INTRODUCTION

The purpose of this study is to develop a theoretical model describing the *working paper review*[1] process in a financial statement audit. The objective of working paper review is to ensure that audit evidence is a sufficient and competent basis for an audit opinion and that audit work has been performed in accordance with generally accepted auditing standards and the firm's own requirements. This review process is one important means of controlling audit quality (Solomon, 1987; AICPA, 1989, Rich, Solomon and Trotman, 1997).

The majority of prior research uses the experimental approach to study the effectiveness and efficiency of different review processes (e.g., Harding and Trotman 1999; Bamber and Ramsay 2000; and Fedor and Ramsay 2007). Our study supplements the literature by providing a theoretical framework for the working paper review process. We show that a formal model which formalizes existing empirical findings sheds light on the choice of the best review process.

Lack of a framework limits the usefulness of the conclusions drawn in some experimental studies. For example, Ramsay (1994) finds that the performance of a senior-senior team does not differ from that of senior-manager team. This finding is counter-intuitive as other studies tend to find that experience leads to improved performance. Also, whereas the performance of working paper review is expected to depend on the technical competence of reviewers, some experimental studies suggest that there is no significant difference in the technical knowledge and problem solving ability between seniors and managers (Bonner and Lewis 1990, Libby and Frederick

[1] Although we use the term working paper review, the concepts discussed also apply to electronic documentation and review of a financial statement audit.

1990, Tan and Libby 1997). Mixed findings such as these possibly can be better understood within a well-formulated theoretical framework such as that developed in this paper.

Our study analyzes whether the sequential, hierarchical working paper review used in practice is an optimal practice. We investigate whether assurance firms benefit from adopting error-specialized review. Together with the empirical findings related to the effectiveness and efficiency of working paper review (e.g., Bamber and Ramsay 2000), the study attempts to facilitate the strategic design of the working paper review process.

Traditionally, the review process is sequential, hierarchical, and iterative. When the working paper preparer, usually a staff member or a senior, is satisfied with the evidence and the corresponding conclusions, the preparer passes the working papers (or its electronic equivalent) to a reviewer, usually a senior or a manager, who checks for errors or inconsistencies and makes comments. The reviewer then decides to either accept the documentation or ask the preparer to perform additional work or reconcile any inconsistencies. However, after interviewing partners from three large, international audit firms and reviewing in-house documents, Rich, Soloman and Trotman (1997) conclude that there are several emerging trends:

1) Elimination of mandatory, multi-layered and detailed review,

2) Movement away from sequential preparer/reviewer interaction with the preparer performing and documenting the work before submitting the working papers for review,

3) Changes in what is documented in audit working papers, i.e. the information being documented in working papers is much more limited, primarily consisting of the preparer's conclusions and references relevant to audit steps, and

4) Additional reviewer testing of preparers' understanding of various audit procedures to decide whether they need to examine the working papers in detail.

Such strategic changes in the working paper review process suggest that auditing firms are seeking more efficient and effective approaches. The benefit-cost analysis derived from our model provides an analytical framework which explains why the traditional team review may not be the optimal choice for auditing firms. Some audit literature suggests that the working paper review should be specialized by limiting the focus of each reviewer to either mechanical errors or conceptual errors (AICPA, 1989; Bacsik and Rizzo, 1983). Auditing firms do not always follow this advice. Bamber and Ramsay (1997, 2000) conclude that error specialized review actually leads to a less accurate and efficient outcome compared to that of all-encompassing review[2]. Their study does not distinguish the consequence of different types of errors and only focus on overall outcome.

Our model sheds light on the reconciliation of the apparent conflict. When outcome reliability is low, firms benefit more from all-encompassing review than from error-specialized review. However, when the reliability is high, the benefits gained from all-encompassing review may be smaller than the additional operational costs incurred.

We build the theoretical model on the concepts of reliability engineering with auditors being modeled as control components (Srivastava 1985; Srivastava and Ward 1983). The review process is modeled as either single component or two sequential components. The output is 'reliability' which represents the portion of the correct information in the output. Thus, an error-free output has reliability of one ("1.0").

In performing simulation analysis, we use two data sets. The parameters are taken where possible from an empirical distribution, or in some cases, on an assumed distribution. The cost analysis is based on the following assumptions. First, the expected litigation cost is a negative exponential function of reliability. Accordingly, the marginal reduction in expected litigation cost is decreasing with the increase of reliability. Second, the operational cost of each component

[2] In error specialized review, a reviewer is instructed to focus on either mechanical errors or conceptual errors. In contrast, a reviewer focuses on both types of errors in all-encompassing review.

is assumed to be constant across different reliability values since the cost of training and employment is fixed for individuals who perform the review process. Using the model, we replicate the previous empirical findings on the performance of individual and team reviews and the outcome of encompassing and error specialized reviews.

The paper is organized as follows. Section II introduces the basic theoretical model. A formula that includes proxies for the determinants of audit performance is derived for a sequential review model with a one- or two-step review process. Section III develops four hypotheses and provides simulation results. Section IV describes two additional propositions. Section V summarizes the conclusions and limitations of the study.

II. RELIABILITY MODELS OF REVIEW PROCESSES

Reliability models were initially developed within engineering and later were applied to internal control evaluation in auditing (Cushing, 1974; Bodnar 1975). However, mechanical system models can not simply be adapted to the information processing domain. For example, Bodnar (1975) found that the more control components in series, the less reliable the system. Intuitively, one would expect that a system with more controls would be more reliable. Bodnar's counter intuitive finding was due to inappropriate application of reliability modeling of controls which involve human components.

Srivastava (1985) points out that a human control component differs from a mechanical component in four ways:

1) A mechanical component either works or does not work, whereas a human component has decision making capability which can lead to differing degrees of failure or success.

2) Unlike a mechanical component which tends to degrade over time, a human component may become more effective and efficient because of learning and experience;

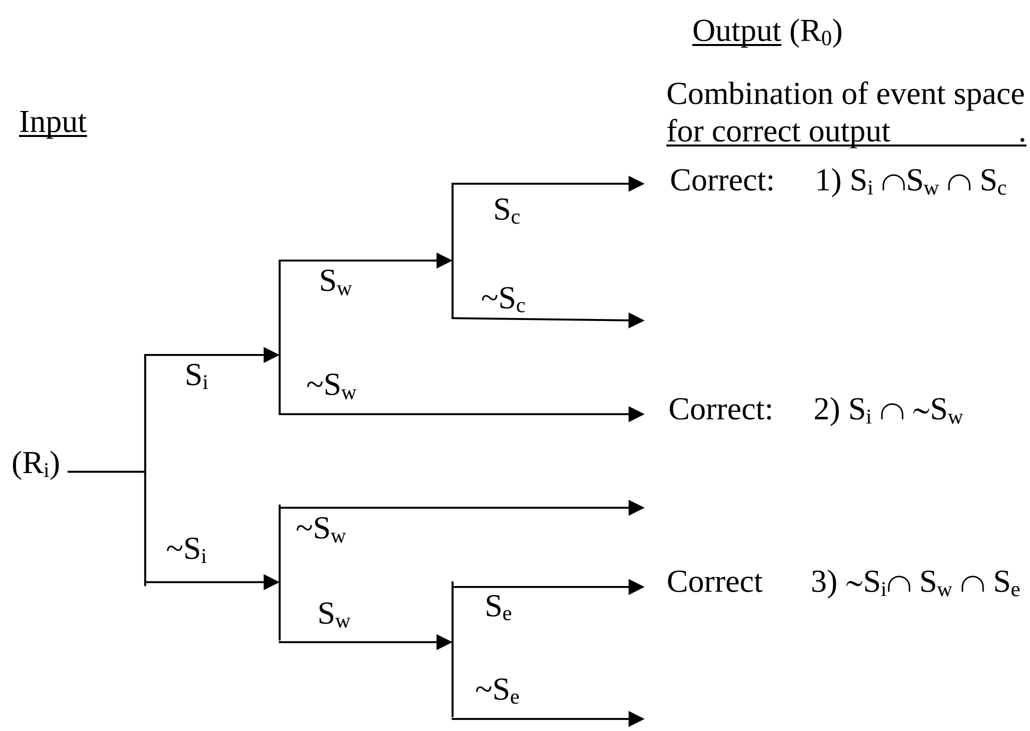
3) The reliability of a mechanical component is independent of input, whereas the reliability of a human information processing component depends on the nature of the input and

4) A mechanical component consists of a single reliability parameter, whereas a human component has multiple parameters.

The most distinguishing features of a human control component are that it has a self-correction function and that it may introduce new errors in the control process. Incorporating new features within the model, Srivastava showed that Bodnar's counter intuitive finding no longer exists under certain circumstances.

Reliability Model of a Review Process with a Human Element

Reliability models of control systems with human elements were first modeled by Srivastava (1982, See also, Srivastava and Ward 1983, and Srivastava 1985). We use Srivastava's approach to model a review process in the context of a financial statement audit. A basic information processing model of working paper review that includes a human component is composed of three branches. The tree structure of the model is as follows.



The symbols in the above diagram are defined in Table 1.

From the above, we can derive the following expression for output reliability in terms of the probability, P, that the output is error free:

$$R_0 = P(S_i \cap S_w \cap S_c) + P(S_i \cap \sim S_w) + P(\sim S_i \cap S_w \cap S_e) \quad (1)$$

Whether the review component works or not is independent of input reliability, we have

$$P(S_i \cap S_w \cap S_c) = P(S_c|S_i \cap S_w)*P(S_i \cap S_w) = P(S_c|S_i \cap S_w)*P(S_i)*P(S_w) = P_c*R_i*P_w$$

$$P(S_i \cap \sim S_w) = R_i*(1 - P_w)$$

$$P(\sim S_i \cap S_w \cap S_e) = P(S_e|\sim S_i \cap S_w)*P(\sim S_i \cap S_w)$$

$$= P(S_e|\sim S_i \cap S_w)*P(\sim S_i)*P(S_w) = P_e*(1-R_i)*P_w$$

Equation (1) is thus rewritten as

$$R_0 = P_c R_i P_w + R_i (1- P_w) + P_e (1- R_i) P_w \quad (2)$$

In this model, the review process is decomposed into three stages. First, the component receives either correct or incorrect information. Second, the component works or does not work. Finally, the component either corrects or is unable to correct errors when the input is incorrect. In this stage human experience, ability, and knowledge is important. If the input is correct at this point, the component either keeps the correct information or makes new errors and thus further contaminates the information. The output reliability R_0 for a single control component is derived above in equation (2).

The three terms in the right hand side of equation (2) stand for three parts of the correct information in the output information. The first two terms are the portion of the correct information in output when the input information is correct. If the human component does not make new errors given correct information, the addition of the two terms equals to the input reliability R_i. The third term is the portion of incorrect information being corrected.

Apart from the assumption that a human component may make errors when it has correct input, two additional assumptions are implied in Srivastava's model. First, P_c and P_e are the same for all control components. Second, P_w, the probability that a control component is in operation, is constant over time. However, both assumptions may not hold in certain audit situations. The values of P_c and P_e reflect the level of an auditor's experience, ability and knowledge, and the value of P_w represents motivation and effort level.

For example, if the senior's and manager's reviews of working papers are recognized as two separated control components, the senior's capability of handling errors may differ from that of a manager. Therefore, P_e and P_c are modeled as different for the two components. In addition, these values may be different even among seniors or managers. Seniors or managers with specialized knowledge about a client's industry have higher values for P_e and P_c than those of other seniors or managers without such knowledge.

The probability that the component is in operation, P_w, may also vary over time and across tasks. For example, when two components work interdependently, reduced motivation or "output interference" may be created whereby hearing what one member says inhibits ability of other members to be able to generate alternative solutions (Hoch, 1984; Owhoso, Messier, and Lynch 2002). Consequently, P_w would change.

Each stage in the model is assumed to be independent from each other. The characteristics of reviewers do not have any control on input reliability. P_e and P_c are independent of P_w as many accounting and psychology studies show that high motivation does not necessarily result in good performance (Bonner and Sprinkle, 2000).

Reliability Model of a Sequential Review Process with Human Elements

The sequential, hierarchical, communication process with two reviewers (assumed to be a senior and manager) is depicted in Figure 1a. The working paper preparer submits working papers to a senior and the senior passes the working papers to a manager after reviewing it. The reviewers may make some notes which will be returned to the preparer to revise the working papers. Both reviews are treated as human control components. The output of the first component is the input of the second component. Such a model is the core of some other more detailed models (Owhoso, et al, 2002; Rich, et al, 1997).

----- Insert Figure 1 here -----

We now derive the formula for the above model based on the concept of a human control component. Seniors and managers might focus on different types of errors in working paper review. For example, some auditing firms require that seniors perform detailed reviews and focus on mechanical errors and that managers perform general reviews and focus on conceptual errors. Seniors and managers may have different detection rates for these two types of errors. These features are incorporated into the formula derived below.

For a single review process case with two types of errors, the output reliability is (see Appendix A and Table 1)

$$R_0 = R_i + (1-R_i)*P_w*[(P_{em} - P_{ec})*R_m + P_{ec}] \quad (3)$$

The difference between this model and the basic model illustrated in Equation (2) is that in this model stage three is further decomposed into two parts – the detection of conceptual and mechanical errors.

For the two control components case, the output reliability is (see Appendix A):

$$R_{02} = R_i + (1-R_i)*P_{w1}*[(P_{em1}-P_{ec1})*R_{m1} + P_{ec1}]$$

$$+ \ (1-R_i)*[1 - P_{w1}*[(P_{em1}-P_{ec1})R_{m1}+ P_{ec1}]]*P_{w2}*[R_{m2}(P_{em2}-P_{ec2})+P_{ec2}] \quad (4)$$

Subscripts 1 and 2 stand for reviewer 1 and 2, respectively. R_{m2} and R_{m1} have the following relationship (see A5):

$$R_{m2} = R_{m1}*(1 – P_{w1}*P_{em1}) / [R_{m1} (1 – P_{w1}*P_{em1}) + (1 - R_{m1})(1 – P_{w1}*P_{ec1})] \quad (5)$$

From equations (4) and (5), we are able to calculate the output reliability, R_{02}, given the parameters R_i, R_{m1}, P_{w1}, P_{w2}, P_{em1}, P_{ec1}, P_{em2}, and P_{ec2}. We will use empirical data to determine these input parameters in section III. Before that, we discuss two extreme cases – reviews with perfect and imperfect error detection. Understanding these two cases helps us see how the outcome is changed when the review process becomes imperfect.

Case 1: Reviews with perfect error detection

First, we assume that two reviewers are instructed to perform an error specialized review, that is the first reviewer focuses on mechanical errors and the second reviewer focuses on conceptual errors (i.e., $P_{ec1} = 0$, $P_{em2} = 0$). We also assume that two reviewers are able to detect each type of errors perfectly ($P_{em1} = 1$, $P_{ec2} = 1$) and that they put full effort ($P_{w1} = P_{w2} = 1$). Substituting these parameter values into equations (4) and (5), we obtain the output reliability R_{02} to be 1. This result makes logical sense because if the first reviewer detects all the mechanical errors and the second reviewer detects all the conceptual errors, then we do expect an error free output.

Second, it is logical for the case where two reviewers are instructed to perform an all-encompassing review instead of an error specialized review. In this case, the output reliability R_{O2} also equals 1. With perfect error detection, all-encompassing review is more efficient because the first reviewer detects all errors and the second reviewer becomes redundant.

However, when the reviewers are not able to correct all errors, i.e. P_{em1}, P_{ec1}, P_{em2}, and P_{ec2} are less than 1, the scenarios become more complex.

Case 2: Reviews with imperfect error detection

In Case 2, we assume that reviewers are only able to detect and correct half of each type of errors, so $P_{w1} = 1$, $P_{w2} = 1$, $P_{em1} = 0.5$, $P_{ec1} = 0$, $P_{em2} = 0$, and $P_{ec2} = 0.5$. We also assume that half of the incorrect information is mechanical and the other half is conceptual, i.e., R_{m1} is equal to 0.5. For the case where reviewers conduct error specialized review, from equation (4) and (5) we calculate that R_0 equals $0.5 + 0.5R_i$. Accordingly, if the reviewers are instructed to perform an all-encompassing review, we have $P_{em1} = P_{ec1} = P_{em2} = P_{ec2} = 0.5$. The relation between R_{O2} and R_i then changes to $R_{O2} = 0.75 + 0.25R_i$.

The result of all-encompassing differs from that of error specialized review. The difference between the outputs in these two cases is $0.25 – 0.25*R_i$. Because R_i is smaller than 1, R_{O2} in all-encompassing review is always bigger than R_{O2} in error specialized review. This implies that all-encompassing review is more effective when both reviewers can only detect half of each type of errors. In general, we can conclude that an all-encompassing review is more effective than an error specialized review under situations of imperfect error detection and full effort.

Reliability Model of a Parallel Review Process

Bamber and Ramsay (1997) examine the combined senior and manager team performance of working paper review, by randomly pairing seniors and managers. The pairing process in their study was not a sequential process, because both managers and seniors had the same working papers to review. Under this circumstance, the senior and manager reviews are parallel. Owhoso et al (1999) call such a parallel review process as “nominal team review” where. both seniors and managers review the working papers without consulting each other's work. This model uses a simple rule, $R_0 = R_{01} + R_{02} - R_{01} * R_{02}$ where, any error caught by one or both reviewers is assumed to be successfully caught by the team. Such a process is depicted in Figure 1b. Such a parallel review is less efficient because one reviewer replicates some work, which has been done by another reviewer, so we only focus on the sequential process here.

Cost Function

Improving the reliability of information is one objective of working paper review. Increasing the reliability of working papers can be expected to increase audit effectiveness. However, overinvestment in the review process reduces audit efficiency. For example, accounting firms can add multiple components and multiple processes to attain very high reliability in working papers, but they would incur tremendous operating cost. The optimal choice in designing a working paper review process involves the tradeoff between effectiveness and efficiency. It is intuitive that marginal benefit decreases with an increase of output reliability. When the reliability is very high, the marginal benefit of adding a control component is expected to be smaller than the cost incurred. The marginal benefit of increasing reliability here is equivalent to the reduction in expected litigation cost due to errors. Here we assume such litigation cost C is a negative exponential function of reliability; that is $C = \alpha * \exp(-\beta R_0)$, where R_0 is the reliability and α and β are the coefficients which are positive. The expected total cost of working paper review is thus

$$E(C_i)= \sum C_i + \alpha * \text{Exp}(-\beta R_0) \quad (7)$$

Where Ci is the operating cost of each control component, i is the number of control components, R_0 is the output reliability derived from previous equations. The first term of the right side of Equation (7) is the total operation cost of control components. The second term is the expected litigation cost caused by imperfect working papers. However, the litigation cost function is not calibrated. This implies that even if working papers are perfect (R_0 =1) there is still an expected litigation cost α*exp(-β). Whereas adding a control component reduces the value of the second term through increasing R_0, the operating cost increases.

III. HYPOTHESES AND SIMULATION RESULTS

As discussed earlier in the introduction, mixed findings of experimental studies can be better understood within a well-formulated theoretical framework such as the one developed in this paper. To understand the mixed results of the experimental studies, we develop several hypotheses that are relevant to the experimental studies, based on our theoretical framework, and test theses hypotheses using simulations.

Before we develop the hypotheses and verify them through simulations, we need to assign values to the input parameters. The parameters R_i and R_m are associated with the characteristics of working paper preparers, such as experience, knowledge, ability and task complexity. P_{w1} and P_{w2} are directly related to the motivation and input efforts of reviewers. P_{em1}, P_{ec1}, P_{em2}, and P_{ec2} are the parameters of reviewers. The fact that these parameters may change simultaneously increases the difficulty and effectiveness of the simulation. We adopt two sets of parameter values in our simulations.

First, we use the empirical data of Ramsay (1994) and Bamber and Ramsay (1997) as the means for determining parameter values (called empirical distributions). Panel A of Table 2 extracts the data from Ramsay (1994) and Bamber and Ramsay (1997). With the assumption that the population of four parameters are normally distributed, the differences of these parameters between Ramsay (1994) and Bamber and Ramsay (1997) are insignificant at the 0.01 level (t =

0.15, 1.39, 0.54, and 0.82 for P_{em1}, P_{em2}, P_{ec1}, P_{ec2}, respectively). These results show that the variance of the detection rates across two studies is small[3].

Based on these observations, we choose the following distributions[4] P_{em1} ~ N (0.69, 0.18), P_{ec1} ~ N(0.65, 0.20), P_{em2} ~ N (0.59, 0.19), P_{ec2} ~ N (0.74, 0.18) and exclude the randomized data which are outside the ranges. The total sample size for each simulation is 1,000. Furthermore, considering that the detection rates of the second reviewer might be influenced by the outcome of the first reviewer, we investigate scenarios where the means of P_{em} and P_{ec} of the second reviewer are reduced by 10%, 20% and 40%.

----- Insert Table 2 here -----

We use two additional hypothetical means of P_{em} and P_{ec} with the standard deviations assumed to be 0.18 for all the detection rates (called hypothetical distributions): (P_{em1}, P_{em2}, P_{ec1}, P_{ec2}) equal to (0.6, 0.4, 0.4, 0.6) or (0.9, 0.7, 0.7, 0.9). These values are used to simulate the beliefs that seniors are better in detecting mechanical errors and managers are superior to seniors in detecting conceptual errors, and that the detection rates of reviewers with specialized industry knowledge are higher than that of reviewers without such knowledge.

After having assumed the above input parameters values, we proceed to generate hypotheses and test them through simulations.

Combined Review versus Individual Review

Auditing firms have adopted hierarchical team review instead of individual review extensively in the past decades. Prior studies show that working paper review process is beneficial and that team review is superior to individual review (Trotman 1985, Trotman and

[3] The range in Ramsay's study is the detection rate for each subject and the range in Bamber and Ramsay's study is for each question.

[4] A normally distributed parameter, P, is represented by P ~ N (μ, σ), where μ is the mean value of the parameter, and σ is the standard deviation.

Yetton 1985). However, Rich et al (1997) find that some auditing firms are moving away from hierarchical review. One potential explanation is that hierarchical review is not an efficient approach even if it is effective. These observations direct us to the following hypothesis.

Hypothesis 1: The performance of combined review is more effective than either a senior or a manager review. However, the increase of the operating cost of a combined review is expected to exceed the reduction in expected litigation cost.

To compare individual review with team review, we compare one control component system with a two-sequential-component system. We use one component to model senior review and two components to model senior-manager or senior-senior team review. Assuming that 80 percent information cues and conclusions in the working papers are correct, i.e. $R_i = 0.8$; that 36 percent of errors are mechanical errors[5], i.e. $R_m = 0.36$; and that both reviewers put in full effort, i.e. $P_{w1} = P_{w2} = 1$, we simulate the output reliability by inputting the empirical distributions of P_{em1}, P_{ec1}, P_{em2}, and P_{ec2}.

The simulation result in Figure 2A supports the first part of the hypothesis. The mean for senior-manager review is 0.977, and the mean for single senior review is 0.93. These means are significantly different (Z= 50.5). The conclusion holds for the comparison between manager review and senior-manager review and it is also supported by the simulations with the hypothetical distributions. The conclusion is fairly intuitive, the second reviewer will probably correct some errors missed by the first reviewer.

To test the second part of hypothesis 1, we need the values of α, β and the operation costs of senior and manager reviews. No data is available for us to estimate α and β in the litigation cost function. We only know that an increase of reliability will reduce expected litigation cost. Thus, we choose hypothetical values for α, β and the operation cost.

---

[5] Bamber and Ramsay seeded 4 mechanical errors and 7 conceptual errors in their case materials, so we choose R_m=4/11 (0.36) in order to make the simulation results comparable.

The purpose of such a simulation is to illustrate how the optimal choice of audit review process could be altered with a change in the litigation cost function, the operation cost of each component, and the number of reviews. When operating expense is small, the result shown in Figure 2A implies that a team review is preferred for auditing firms because the higher reliability would lead to a significant reduction in expected litigation cost.

However, when expected litigation cost is at the same order of operating cost, the reliability output based on the empirical distributions may leads to an opposite conclusion. With α=1000 and β=10, the expected litigation cost is in the range of [0.045, 0.203] for the reliability R_0 = [0.85, 1.0]. Figure 2B presents three scenarios in which the operating cost of a senior review is 0.12 (S) and the operating cost of a senior-manager review is 0.13 (S-M1) and 0.18 (S-M2), respectively.

Simulation results suggest that the total cost for a senior review could be higher or lower than a team review, depending on the reduction of litigation costs and the incremental operating cost of an additional reviewer. When the operating cost of a senior review and of a senior-manager review is 0.12 and 0.18, respectively, the mean value of total cost for senior review is lower than team review. In this case the optimal choice is senior review rather than senior-manager team review.

----- Insert Figure 2 here -----

Senior-manager Team Review versus Senior-senior Team Review

Audit judgment performance is affected by an auditor's experience, ability, knowledge, and motivation (Libby and Luft 1993, Bonner and Pennington 1991). An auditor is not able to obtain experience and knowledge in short time. They are accumulated through many different audit engagements and over time. Usually, managers have more knowledge and experience than seniors though it is possible that the difference in technical knowledge and problem solving ability between seniors and managers is not significant. Logically, a senior-manager team should be more effective than a senior-senior team in reviewing working paper. On the other hand, a senior-manager would incur more operation cost. Such observations lead to our second hypothesis.

Hypothesis 2: Review performance of the team consisting of an audit senior and an audit manager is expected to be better than that of the team consisting of two seniors. The expected cost of a senior-manager team review is expected to be higher than that of a senior-senior team review.

The first simulation is based on our hypothetical distributions, i.e., (P_{em1}, P_{em2}, P_{ec1}, P_{ec2}) equal to (0.6, 0.4, 0.4, 0.6) or (0.9, 0.7, 0.7, 0.9) We find that the performance difference between two types of teams is significant if the detection rates for a senior are low and those for a manager are high. The difference becomes smaller when the detection rates for both the senior and the manager are close to each other. These results support the first part of the hypothesis.

We repeated our simulation with empirical distributions. R_i, R_m, P_{w1} and P_{w2} are the same as those used in testing hypothesis 1. The mean of output reliability for senior-senior pair review is 0.977 and for senior-manager pair review is 0.976. The difference is not significant at 0.01 level. The empirical and hypothetical distributions lead to opposite conclusions. The reason is that the error detection rates for a senior and a manager in the empirical distributions are very close to each other.

The result based on empirical distributions is consistent with several prior studies: The performance of a review team consisting of a senior and a manager is not different from the performance of a review team consisting of two seniors (Trotman and Yetton 1985). Only small difference in technical knowledge results from experience beyond the senior level (Bonner and Lewis 1990, Libby and Frederick 1990).

These findings are significance for practice. If there is no difference, an auditing firm would be better off by adopting senior-senior teams. While the cost is reduced, the quality of the working paper review remains unchanged.

To test the generality of these results, we perform a sensitivity analysis. If the senior's detection rate for both mechanical and conceptual errors drops 20% from the original empirical means, the output reliabilities for the two types of teams are different at 0.01 level. Thus, evidently, there are critical parameter values for the hypothesis to be valid. We find that improving the detection rates of a senior has the same effect on the output reliability as improving the manager's detection rate. This implies that training a senior to do a better job is a more efficient approach. The conclusion also implies that we may see the difference in the performance of working paper review between senior-senior and senior-manager teams if these teams do not have enough specialized experience and knowledge on a client's industry.

However, the second half of the hypothesis may not be supported if the expected litigation cost functions resulted from two types of team reviews are different. The previous argument focuses on the overall reliability and ignores the potential change in the litigation cost. Though there is no difference in the overall outcome reliability distribution, the expected litigation cost function for the two types of team reviews would be different. The empirical distribution indicates that, in general, managers are better in detecting conceptual errors. The uncorrected portion of the output in a senior-manager team review has few conceptual errors than that in a senior-senior team review. Prior study shows that litigation is more likely to occur when auditors fail to detect commonly occurring frauds (Bonner, Palmrose and Young 1998). It is reasonable to assume that few conceptual errors left in the output will result in the reduction of expected litigation cost. Consequently, the parameter α is different for two types of team reviews. Simulations indicate that the conclusion drawn from benefit-cost analysis is sensitive to the change of the value of α, but senior-manager team review is more likely to be preferred because the reduction in expected litigation cost exceeds the increase in operating cost.

Senior-manager Team with Specialized Industry Knowledge versus Senior-manager without Specialized Industry Knowledge

Prior research has shown that knowledge about a special industry will affect review performance in audit engagements (Owhoso et al. 2002). Auditors with specialized industry knowledge and experience find and correct errors that are rarely found by auditors without such knowledge and work experience. The parameters for these two types of auditors would be significantly different. Accordingly, we generate our third hypothesis.

Hypothesis 3: The performance of a senior-manager team with specialized industry knowledge is expected to be better than a senior-manager team without such knowledge. The cost of the former is thus expected to be smaller than that of the later.

We assume that the team with specialized knowledge (team 1) has the mean values of P_{em1}, P_{em2}, P_{ec1}, and P_{ec2} as 0.9, 0.7, 0.7, and 0.9, respectively, and that the team without specialized knowledge (team 2) has the values 0.6, 0.4, 0.4, and 0.6, respectively. Considering that both teams have the same incentive to perform well, we choose that P_{w1} and P_{w2} equal to1. R_m and the standard deviations for P_{em1}, P_{em2}, P_{ec1}, P_{ec2} are assumed to be the same as used in the previous simulations.

Figure 3 compares the results of team 1 with that of team 2. The mean of output reliability for team 1 is 0.988, and the mean for team 2 is 0.952. They are significantly different (Z= 55.69).

The simulation results shown in Figure 3 support hypothesis 3. Empirical studies in auditing also suggest that Big 5 auditing firms with specialized industry expertise have lower frequencies of getting involved with litigation (Palmrose 1988). While the operating cost is equivalent, the expected litigation cost is smaller for an industry specialized review. This confirms that the investment in training auditors to get industry knowledge is an efficient and effective approach to reach higher quality working paper review.

----- Insert Figure 3 here -----

All-encompassing Review versus Error Specialized Review

We do not have enough evidence in the literature to show that error specialized review is better than all-encompassing review. However, in the prior analysis of special case 2, we show that, in general, an all-encompassing review is more effective than an error specialized review under the assumption of imperfect error detection. Error specialized review is perceived to be a more efficient approach because of less effort required, but Bamber and Ramsay (2000) find the opposite results. We thus hypothesize that there is a performance difference between the two types of reviews - all-encompassing review and error specialized review.

Hypothesis 4: Ceteris Paribus, the performance of an all-encompassing review is expected to differ from the performance of an error specialized review.

The simulation with empirical distributions supports the hypothesis and shows that all-encompassing review is better than error specialized review given the constant P_w. The mean of output reliability for all-encompassing review is 0.975 and the mean for specialized review is 0.970, Z value is 8.55. The conclusion, though should be interpreted cautiously because of the small difference between two values. This result is consistent with the empirical findings of Bamber and Ramsay (1997) and Harding and Trotman)1999) that all-encompassing review teams are more effective than error specialized review teams.

All previous simulations are based on the assumption that the second reviewer is as effective as the first one. This assumption may affect the validity of the simulation results. In fact, if the first reviewer has corrected the errors easily to be detected, the second reviewer will have a harder time in detecting the remaining errors and hence the detection rates of the second reviewer may decrease. The parameter P_w for the second reviewer would also change. The parameters for two reviewers are thus not independent.

Panel B of Table 2 showing the data extracted from the previous study (Bamber and Ramsay, 1997) indicates that there is a significant overlap between errors detected by the first and second reviewer. Table 3 shows that the simulation results for hypotheses 1 and 2 hold when the means of P_{em2} and P_{ec2} are reduced by 10%, 20% and 40%. We also test hypothesis 3 but not hypothesis 4 when P_{em2} and P_{ec2} are reduced by 10% and 20%. In hypothesis 4, the second reviewer does not have a reduction of detection rates in an error specialized review. This is because, under error specialized review, each reviewer detects one type of error.

----- Insert Table 3 here -----

IV. ADDITIONAL PROPOSITIONS

Proposition 1: *The parameters representing motivation and effort level, P_{w1} and P_{w2}, for the two reviewers are not of the same importance in the working paper review process.*

Several parameters in our model are proxies for knowledge, experience, ability, and motivation. These are factors studied frequently in the audit judgment research (Libby and Luft, 1993). Both seniors and managers are capable of detecting mechanical and conceptual errors in all-encompassing review, but the motivation and effort levels represented by P_{w1} and P_{w2} may be different. If a senior and a manager have different capability in finding and correcting errors, their incremental contributions to the output reliability should be different.

Derived from equation (4), we have partial derivatives of R_0 with respect to P_{w1} and P_{w2} as:

$$\partial R_0/\partial P_{w1} = (1- R_i)*K_1*(1-P_{w2}*K_2)$$

$\partial R_0/\partial P_{w2} = (1- R_i)*K_2*(1-P_{w1}*K_1)$

where $K_1 = (P_{em1}-P_{ec1})*R_{m1} + P_{ec1}$, $K_2 = (P_{em2} - P_{ec2}) * R_{m2} + P_{ec2}$. Therefore,

$$(\partial R_0/\partial P_{w1}) / (\partial R_0/\partial P_{w2}) = (K_1 - K_1*K_2*P_{w2})/ (K_2 - K_1*K_2*P_{w1}) \quad (10)$$

If $K_1 > K_2$, and $P_{w1} = P_{w2}$, from equation (8),

$(\partial R_0/\partial P_{w1}) / (\partial R_0/\partial P_{w2}) > 1$

The above condition implies that the change in output reliability is larger when P_{w1} changes than that when P_{w2} changes. In contrast, if $K_1 < K_2$, the change in output reliability is smaller when P_{w1} changes than that when P_{w2} changes. Therefore, P_{w1} and P_{w2} are not of the same importance in working paper review process.

Furthermore, let us use a numerical example to show how these input parameters affect the output reliability. Substituting the mean of P_{em1}, P_{ec1}, P_{em2}, P_{ec2} into equation (5), and setting $R_i = 0.8$ and $R_m = 0.36$, we have output reliability

$$R_{02} = 0.8 + 0.1255P_{w1} + 0.1362P_{w2} – 0.0855P_{w1}*P_{w2} \quad (11)$$

The difference between the coefficients of P_{w1} and P_{w2} supports the proposition. In equation (11), we have assumed that P_{w1} is independent of P_{w2}. This assumption becomes invalid when the second reviewer increases his/her effort, P_{w2}, after knowing the fact that the first reviewer has put less effort (small P_{w1}).

Proposition 2: *The increase in the quality of working paper input will not automatically increase the reliability of working paper output.*

In a mechanical control system, the output reliability always increases with the input reliability. However, in a control system with a human component, this is not necessarily true. If the human component perceives the change in input quality, they may adjust their control parameters accordingly. In the working paper review process, if a reviewer has trust in the preparer, the reviewer may think the input quality is high, i.e. R_i is close to 1. The reviewer thus may reduce their effort in reviewing the papers, i.e. P_w becomes smaller.

Derived from equation (3), the partial derivative of R_{02} with respect to R_i is:

$$\partial R_{02}/\partial R_i = 1+ (1-R_i)*\partial P_w/\partial R_i* K – P_w*K, \quad (12)$$

where $K = [(P_{em}-P_{ec})*_{Rm} +P_{ec})]$. The sign of $\partial P_w/\partial R_i$ would logically be negative or zero. $\partial P_w/\partial R_i$ is considered to be zero when a reviewer's effort is independent of the input reliability. For this case, $\partial R_0/\partial R_i$ = 1- P_w*K, which is always positive, implying that R_{02} will increase with the increase in R_i irrespective of the value of P_w. However, when a reviewer perceives that the reliability of working papers increases, he may reduce his efforts (P_w) in review. For this situation, the sign of $\partial P_w/\partial R_i$ is thus negative. This condition may lead to a negative value for $\partial R_{02}/\partial R_i$. The negative sign of the derivative, $\partial R_{02}/\partial R_i$, implies that the increase in the quality of working paper input will not necessarily increase the reliability of working paper review if the reviewer reduces his/her effort level with the increase of input reliability.

This finding is consistent with empirical findings. Asare and McDaniel (1996) find that reviewers of unfamiliar preparers re-perform more of the preparers' work. They also find that task complexity and preparer familiarity[6] together determines the effectiveness of detecting errors. Reviewers of familiar preparers are more effective on the complex tasks whereas reviewers of unfamiliar preparers are more effective on the routine tasks.

V. CONCLUSIONS AND LIMITATIONS

This study provides a basic theoretical model for the audit working paper review process and evaluates important empirical findings through simulations and differential analysis. The most important findings include the following.

[6] Determined by whether the reviewer is familiar with the preparer.

First, the performance of team review is more effective than either senior or manager review. Second, there is no difference in review performance between a senior-manager team and a team consisting of two seniors if seniors and managers work at the same level of effectiveness. Third, the performance of a senior-manager team with specialized industry knowledge is better than a senior-manager team without such knowledge. Finally, there is no performance difference between an encompassing review and an error specialized review when the error detection rates are high.

Whereas these findings in the effectiveness dimension of working paper review are fairly intuitive, the simulation and analytical results provide new insights on efficiency. Although team review increases the reliability of working papers, it may not be an optimal choice when the expected litigation cost is at the same order of magnitude as the operating cost of each review. The increase in the operation cost may exceed the reduction of the expected litigation cost when the reliability is already high. Though empirical findings show that the difference in technical knowledge between seniors and managers is often insignificant and that the performance of a senior-senior review does not differ from that of a senior-manager review, it is not a simple matter to replace a senior-manager team with a senior-senior team. The expectation that managers will detect more conceptual errors will lead to lower expected litigation cost even if the overall reliability from the reviews of two types of teams are identical.

The study also provides analytical explanations as to why the parameters representing motivation and effort level, P_{w1} and P_{w2} are not of the same importance in the working paper review process. It also shows that an increase in the quality of working paper input does not necessarily increase the reliability of the working paper output if we introduce a motivation variable, a factor affecting audit performance. It provides implications to practitioners as to when they can optimize the role of incentives in the working paper review process.

These findings should be considered along with the study limitations. One limitation is that all detection rates are assumed to be normally distributed, which may not be the case in practice. However, other typical distributions are not expected to affect the simulation results significantly. Using a uniform distribution for the input parameters, we find that all four hypotheses are supported.

Another limitation relates to the assumed values of the input parameters. Our results are sensitive to the mean and variance of these input parameters. If more empirical data were available, we could test the robustness of the model to these values. If a precise litigation cost function can be estimated, auditing firms can benefit from the theoretical model provided in the paper when they make strategic decisions in choosing a specific working paper review process.

REFERENCES

American Institute of Certified Public Accountants (AICPA), 1989, Working Papers Review.

Asare, S. and L. McDaniel, 1996. The effects of familiarity with the preparer and task complexity on the effectiveness of the audit review process. *The Accounting Review*, 71, (April): pp139-60.

Bacsik, J. and S. Rizzo 1983. Review of audit working papers, *The CPA Journal* (November), pp12-20.

Bamber, E.M. and L.S. Bamber 1988. A descriptive study of audit managers' working paper review. *Auditing: A Journal of Practice & Theory* (Spring), pp137-49.

Bamber, E. M. and R. J. Ramsay, 1997. An investigation of the effects of specialization in audit workpaper review. *Contemporary Accounting Research.* 14 (Fall), pp501-13.

Bamber, M. and R. Ramsay, 2000. The effects of specialization in audit workpaper review on review efficiency and reviewers’ confidence. *Auditing: A Journal of Practice & Theory (Fall)*, pp147-57.

Bonner, S., and B. Lewis, 1990. Determinants of auditor expertise, *Journal of Accounting Research*, (supplement), pp1-20.

Bonner, S., Z. Palmrose, and S. Young, 1998. Fraud type and auditor litigation: An analysis of SEC accounting and auditing enforcement releases, *The Accounting Review*, 73, 4, pp 503-32.

Bonner, S., and N. Pennington 1991. Cognitive processes and knowledge as determinants of audit expertise, *Journal of Accounting Literature*, 10, pp1-50.

Bonner, S., and G. Sprinkle, 2000. Propositions about the Effects of performance-contingent monetary incentives on task performance, Working Paper, University of Southern California.

Bodnar, G. 1975. Reliability modeling of internal control systems, *The Accounting Review*, (October), pp747-57.

Cushing, B. 1974. A mathematical approach to the analysis and design of internal control systems. *The Accounting Review*, (January), pp24-41.

Fedor, D. B. and R. J. Ramsay. 2007. Effects of supervisor power on preparers' responses to audit review*: A* field study*, Behavioral Research in Accounting*, 2007, v. 19, pp. 91-105

Harding, N., and K. Trotman, 1999. Hierarchical differences in audit workpaper review performance. *Contemporary Accounting Research*, 16, 4, pp671-84.

Hoch, S. J. 1984. Availability and interference in predictive judgment, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 4, pp649-62.

Libby, R., 1995. The role of knowledge and memory in audit judgement. *In Judgement and Decision Making Studies in Accounting and Auditing*, Ed. Ashton and Ashton, pp176-205.

Libby, R., and D. Frederick, 1990. Expertise and the ability to explain audit findings, *Journal of Accounting Research*, (Autumn), pp348-67

Libby, R., and Luft, J. 1993. Determinants of judgement performance in accounting settings: ability, knowledge, motivation and environment. *Accounting, Organizations and Society*, 18, pp701-16.

Miller, C L.; D. B. Fedor, and R. J. Ramsay 2006. Effects of Discussion of Audit Reviews on Auditors' Motivation and Performance. *Behavioral Research in Accounting*, v. 18, pp. 135-46

Owhoso, V.E., W.F. Messier, Jr., and J.G. Lynch, Jr., 2002. Error Detection by Industry-Specialized Teams during Sequential Audit Review. *Journal of Accounting Research*. 40, 3: 883-900.

Palmrose, Z. 1988. An analysis of auditor litigation and audit service quality, *The Accounting Review*, 64, pp683-709.

Ramsay, R. J. 1994. Senior/manager difference in audit workpaper review performance. *Journal of Accounting Research.* 32, 1, pp127-35.

Rich, J. S., I. Solomon, and K. T. Trotman, 1997. Multi-auditor judgement/decision making research: A decade later, *Journal of Accounting Literature,* 16, pp86-122.

Rich, J.S., I. Solomon and K.T.Trotman,1997 The audit review process: a characterization from the persuasion perspective, Accounting, organizations and society, vol. 22 no. 5, pp. 481-505.

Solomon, I. 1987. Multi-auditor judgment/decision making research. *Journal of Accounting Literature*, 6, 1-25.

Srivastava, R. P. 1982. The Reliability of Internal Accounting Control Systems: Design and Analysis. Ph.D. Dissertation. University of Oklahoma.

Srivastava, R. 1985. A note on internal control systems with control components in series. *The Accounting Review,* (July), pp505-7.

Srivastava, R., and B.H. Ward. 1983. Reliability modeling of information systems with human elements: A new perspective, *IEEE Transactions: Total Systems Reliability Symposium* (December), pp30-9.

Tan, H., and R. Libby, 1997. Tacit managerial versus technical knowledge as determinants of audit expertise in the field, *Journal of Accounting Research*, 35, 1, pp97-113

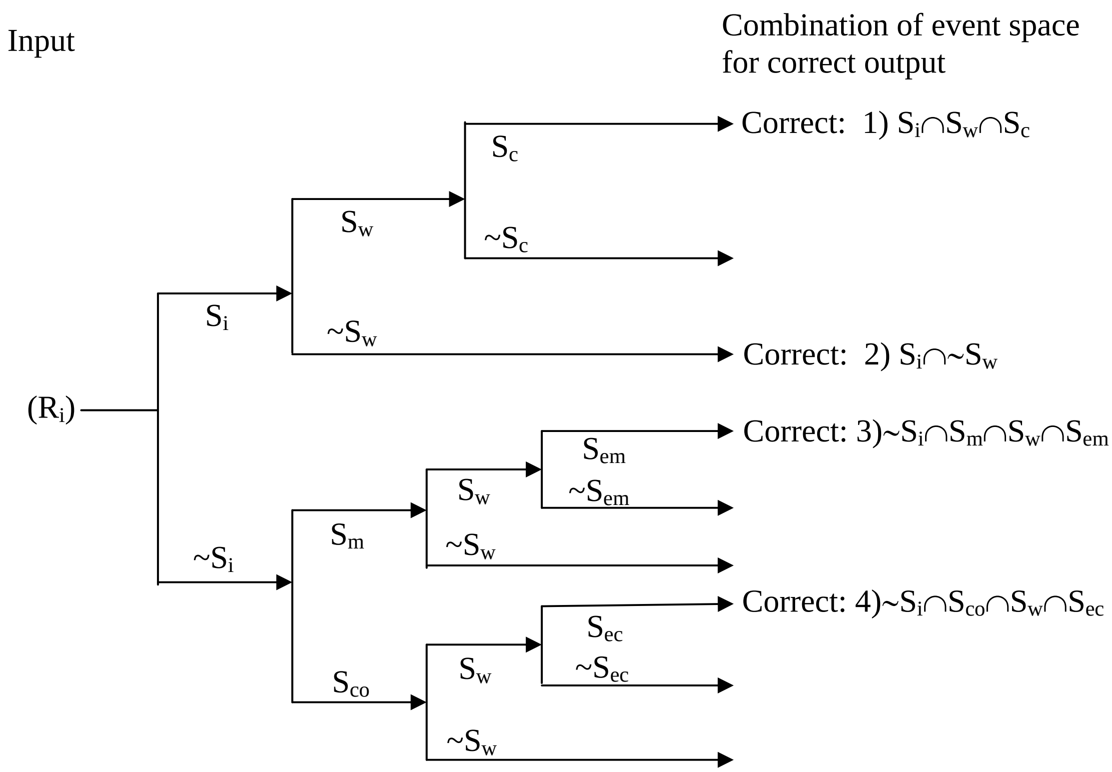
Trotman, K.T. 1985. The review process and the accuracy of auditor judgement. *Journal of Accounting Research.* (Autumn), pp740-52.

Trotman, K.T., and P. W. Yetton 1985. The effect of the review process on auditor judgements. *Journal of Accounting Research*. (Spring), pp256-67.

APPENDIX A

Reliability Model of a Review Process with Single Component and Two Types of Errors: Mechanical and Conceptual

First, we examine a single component case with two types of errors. The following Figure shows the tree structure of such a control component.



The above symbols are defined in Table 1.

Based on the above tree structure, we can derive the output reliability, R_0, as:

$$R_0 = P(S_i \cap S_w \cap S_c) + P(S_i \cap \sim S_w) + P(\sim S_i \cap S_m \cap S_w \cap S_{em}) + P(\sim S_i \cap S_{co} \cap S_w \cap S_{ec}) \quad \text{(A1)}$$

Besides the assumptions in Srivastava (1985), we further assume that reviewers rarely make further errors given correct input information, therefore, $P(\sim S_c|S_i \cap S_w) \approx 0$, and $P(S_c|S_i \cap S_w) \approx 1.0$. Because whether control component can correct errors is not related to whether there are mechanical and conceptual errors in working papers, S_{ec} and S_{em} are independent of S_{co} and S_m, respectively. The equation (A1) is thus simplified to

$$R_0 = P(S_c|S_i \cap S_w)*P(S_i \cap S_w) + P(S_i \cap \sim S_w) + P(S_{em}|\sim S_i \cap S_m \cap S_w)*P(\sim S_i \cap S_m \cap S_w)$$

$$+ P(S_{ec}|\sim S_i \cap S_{co} \cap S_w) * P(\sim S_i \cap S_{co} \cap S_w)$$

$$= P(S_i) + P(\sim S_i)*P(S_m|\sim S_i)*P(S_w)*P_{em} + P(\sim S_i)*P(S_{co}|\sim S_i)*P(S_w)*P_{ec}$$

$$= R_i + (1-R_i)*R_m*P_w*P_{em} + (1-R_i)(1-R_m)*P_w*P_{ec}$$

i.e.

$$R_0 = R_i + (1-R_i)*P_w*[(P_{em} - P_{ec})*R_m + P_{ec}] \quad \text{(A2)}$$

R_m is defined as the probability of mechanical errors given there are errors, i.e., $R_m = P(S_m|\sim S_i)$. It represents the fraction of mechanical errors to total errors. To simplify our discussion, we assume that there is no overlap between mechanical and conceptual errors. This assumption is actually implied in the previous empirical studies.

For the model with two sequential control components, the input of the second component is the output of the first component, $R_{i2} = R_{01}$, therefore, the output reliability, R_{02}, is given by

$$R_{02} = R_{01} + (1- R_{01})*P_{w2} [R_{m2}*(P_{em2} - P_{ec2}) + P_{ec2}]$$

$$= R_i + (1- R_i)*P_{w1}*[(P_{em1}- P_{ec1})*R_{m1} + P_{ec1}]$$

$$+ (1- R_i)*[1 - P_{w1}*[(P_{em1}- P_{ec1})*R_{m1}+ P_{ec1}]]*P_{w2}*[R_{m2}(P_{em2} - P_{ec2}) + P_{ec2}] \quad \text{(A3)}$$

Subscript 1 and 2 stand for reviewer 1 and 2, respectively.

R_{m2} is related to R_{m1}. This relationship can be determined as follows. The portion of errors in the output from the first component is

$$1 – R_{01} = (1-R_i) [1 - R_{m1}*P_{w1}*P_{em1} - (1-R_{m1})*P_{w1}*P_{ec1}]$$

$$= (1-R_i) [R_{m1}(1-P_{w1}*P_{em1}) + (1-R_{m1})(1 - P_{w1}*P_{ec1})] \quad \text{(A4)}$$

The first term on the right hand side of (A4) relates to the mechanical errors and the second term relates to the conceptual errors. Therefore,

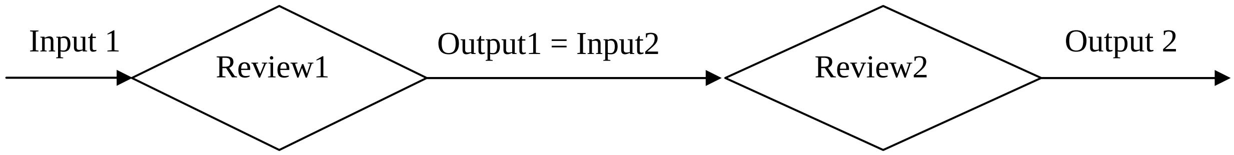
R_{m2} = Uncorrected mechanical errors after review 1 / Total uncorrected errors after review 1

$$= R_{m1} * (1 - P_{w1} * P_{em1}) / [R_{m1} (1 - P_{w1} * P_{em1}) + (1 - R_{m1})(1 - P_{w1} * P_{ec1})] \qquad (A5)$$

With equation (A3) and (A5), we are able to calculate output reliability given the parameters R_i, R_{m1}, P_{w1}, P_{w2}, P_{em1}, P_{ec1}, P_{em2}, and P_{ec2}.

Figure 1: Sequential vs. Parallel Two-Component Review Process.

a) Sequential model:



b) Parallel model:

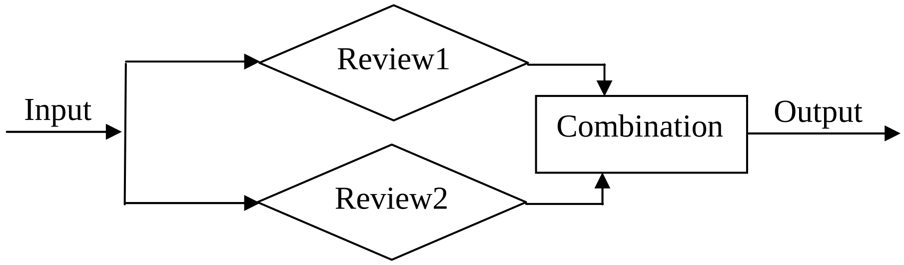
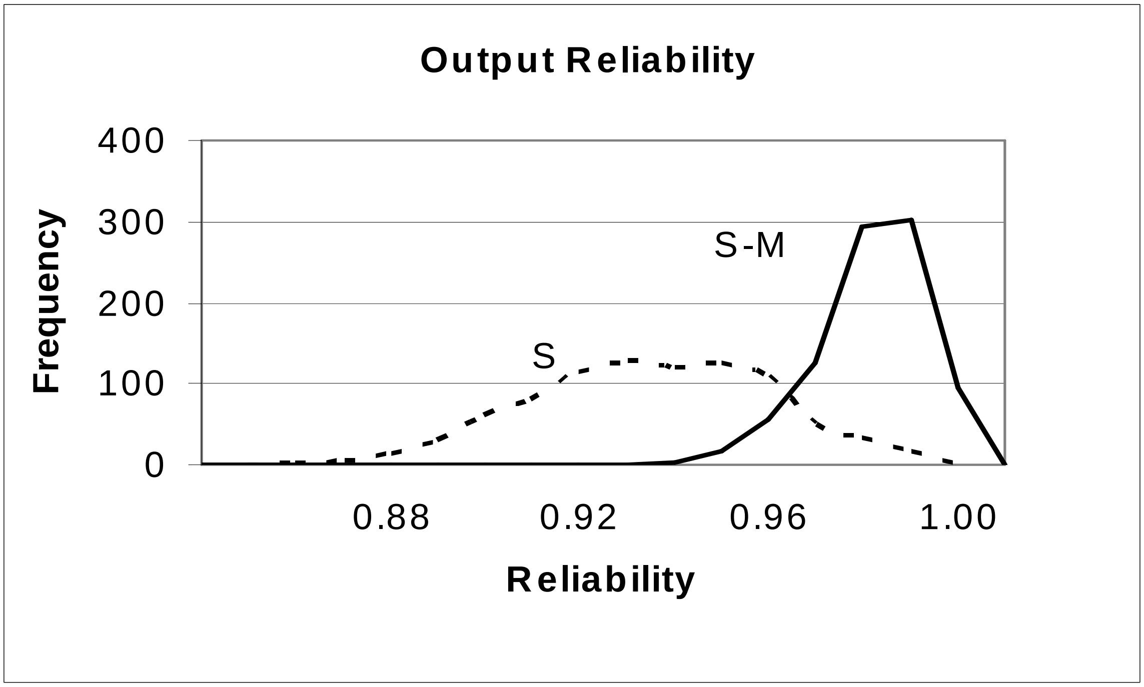


Figure 2

Panel A. The Distributions of Output Reliability of a Senior Review (S) and a Senior-Manager (S-M) Team Review



Panel B. The distributions of the total cost with litigation cost parameters, α = 1000, β = 10, for a senior review (S) with operation cost 0.12, a senior-manager review (S-M1) with operation cost 0.13, and a senior-manager review (S-M2) with operation cost 0.18.

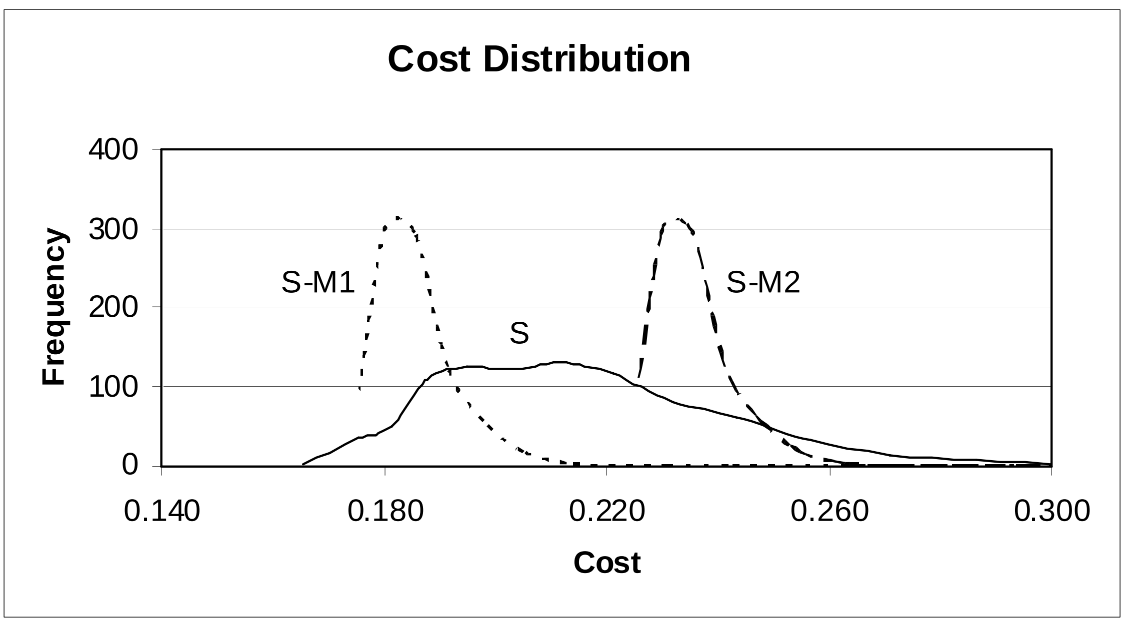


Figure 3

The Comparison of Output Reliability of Two Senior-Manager Review Teams (Team 1- with specialized industry knowledge, Team 2 – without industry knowledge)

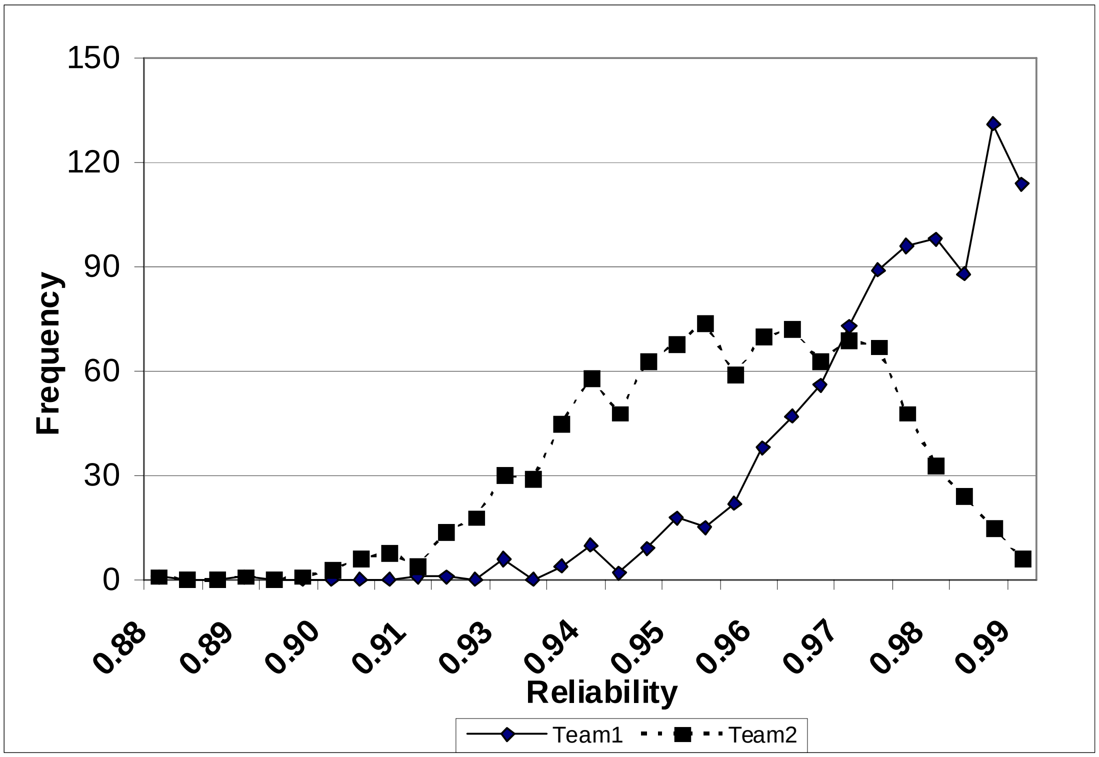


Table 1: List of Symbols and Their Definitions

Symbol	Definition
R_i	Input Reliability
R_O	Output Reliability.
R_m, R_c	The ratios of mechanical errors and conceptual errors to total errors, respectively.
S_i	The state that the input information is correct.
$\sim S_i$	The state that the input information is incorrect.
S_w	The state that the control component operates.
$\sim S_w$	The state that the control component does not operate.
S_c	The state that the control component does not introduce new errors.
$\sim S_c$	The state that the control component introduces new errors.
S_e	The state that the control component corrects input errors.
$\sim S_e$	The state that the control component is unable to correct input errors.
S_m, S_{co}	The state that mechanical and conceptual errors are present, respectively.
S_{em}, $\sim S_{em}$	The state that the control component corrects or does not correct mechanical errors.
S_{ec}, $\sim S_{ec}$	The state that control component either corrects or does not correct conceptual errors.
P_c	$P_c = P(S_c \| S_i \cap S_w)$, the probability that the control correctly processes information given the correct input and the control is in operation.
P_e	$P_e = P(S_e \| \sim S_i \cap S_w)$, the probability that the control corrects errors given that the input information has errors and the control is in operation.
P_w	The probability that a control component is in operation.
P_{wi}	The probability that ith control component is in operation, where i = 1, 2.
P_{em}	$P_{em} = P(S_{em} \| \sim S_i \cap S_m \cap S_w)$, the probability of S_{em} given $\sim S_i$, S_m, and S_w.
P_{ec}	$P_{ec} = P(S_{ec} \| \sim S_i \cap S_{co} \cap S_w)$, the probability of S_{ec} given $\sim S_i$, S_{co}, and S_w.

Table 2: The extracted empirical data of the parameters

Panel A: The Extracted Empirical Data of P_{em1}, P_{ec1}, P_{em2}, and P_{ec2}

	Ramsay (1994)		Bamber and Ramsay (1997)	
	Senior	**Manager**	**Senior**	**Manager**
Mechanical Errors	P_{em1}	P_{em2}	P_{em1}	P_{em2}
Mean	0.694	0.590	0.700	0.50
Std. Deviation	0.179	0.186	0.123	0.261
Range	0.375 -1.0	0.25-0.875	0.55-0.80	0.333-0.889
Conceptual Errors	P_{ec1}	P_{ec2}	P_{ec1}	P_{ec2}
Mean	0.650	0.736	0.679	0.77
Std. Deviation	0.205	0.176	0.196	0.138
Range	0.125 –1.0	0.375-1.0	0.35-0.80	0.50 – 0.89
Sample Size	**20**	**18**	**39**	**35**

Panel B: Errors Detected by Both Seniors and Managers (Source: Bamber and Ramsay 1997)

	Mechanical errors detected by both Seniors and Managers (Mean)*	**Conceptual errors detected by both Seniors and Managers (Mean)***
Error specialized review	0.559 + 0.579 – 0.795 = 0.343	0.664 + 0.632 – 0.819 = 0.477
All encompassing review	0.694 + 0.59 – 0.828 = 0.456	0.65 + 0.736 – 0.883 = 0.503

*Percent of errors detected by both Seniors and Managers = Percent of errors detected by Seniors + Percent of errors detected by Managers – Percent of Errors detected by paired seniors and Managers.

Table 3: The Comparison between the Scenarios with Different P_{em2} and P_{ec2}

	Scenario 1 (10% reduction in P_{em2}, P_{ec2})	**Scenario 2** (20% reduction in P_{em2}, P_{ec2})	**Scenario 3** (40% reduction in P_{em2}, P_{ec2})
Conclusion for H1	Hold (z = 45.0)	Hold (z = 39.4)	Hold (z = 28.3)
Conclusion for H2	Hold (z = 35.9)	Hold (z = 31.3)	Hold (z = 22.5)