Determinants of Audit Preparers’ Workpaper Justifications

Premila Gowri Shankar
Nanyang Technological University

Hun-Tong Tan
Nanyang Technological University

ABSTRACT: Research indicates that justifiability of a decision is a key determinant of work quality. In this paper, we investigate whether and how the requirement for preparers to justify an audit conclusion to a reviewer with task preferences that are similar to or dissimilar from their explicitly stated initial task preferences determine the nature and extent of justifications documented in their workpapers, and how this effect is moderated by their technical and tacit managerial knowledge. Results of an experiment with auditors as participants indicate that auditors use different forms of justifications under different circumstances. Auditors with high tacit managerial knowledge and high technical knowledge have a higher proclivity to enumerate more pros versus cons reasons, and to consider a wider breadth of issues when they are required to justify their conclusions to reviewers with dissimilar task preferences. In contrast, high tacit managerial knowledge auditors required to justify to reviewers with similar initial task preferences have a greater tendency to employ an evidence framing approach, in which language is used to emphasize consistent evidence by downplaying associated inconsistent evidence. In additional analyses, we find that these three specific forms of justifications are positively associated with evaluations of work quality.

Keywords: justification; tacit managerial knowledge; review process.

Data Availability: Contact the authors.

1. INTRODUCTION

The need to justify a conclusion or point of view is a ubiquitous feature in many business and professional contexts. In such settings, the justifier generally seeks to provide support for an expressed viewpoint, and to persuade recipients of its validity. For example, managers routinely justify their assessments of the company’s past and future outlook, and may adopt persuasive strategies in justifying their assessments to investors and analysts (Sedor 2002; Kuperman 2003). Financial analysts also provide justifications for their research reports, and investors’ reliance on these reports is determined in part by the persuasiveness of these justifications (Hirst et al. 1995).

We appreciate the helpful comments of Amit Das, Jun Han, Karim Jamal, Terence Ng, Pearl Tan, Ken Trotman, workshop participants at Nanyang Technological University, Jane Kennedy (editor), and two anonymous reviewers. We thank the auditors from the two Big 4 public accounting firms for participating in our study.

Editor’s note: This paper was accepted by S. Jane Kennedy, Editor.

Submitted June 2004
Accepted August 2005
In this paper, we focus on justifications in an auditing context, where the quality of documented justifications is especially important. Auditing standards mandate that workpapers contain sufficient and appropriate justifications that support a documented conclusion (SAS No. 96, AICPA 2002). Furthermore, such justifications are referred to as a source of evidence by others such as judges, jurors, expert witnesses, and regulators in litigation cases (Kennedy et al. 1997; Beus 2003; Ricchiute 2004). In particular, for ambiguous unstructured audit tasks with no one single appropriate solution, justifiability or defensibility of the documented judgment or decision is a major measure of work quality (Waller and Felix 1984; Emby and Gibbins 1988; Kennedy et al. 1997).

Despite its importance, few studies examine the nature and determinants of justifications documented by workpaper preparers. Two studies are relevant. Koonce et al. (1995) and Tan et al. (1997) provide evidence suggesting that preparers increase the quantity of justifications when they are made accountable to a reviewer with unknown preferences. However, both papers document effects on the quantity not the nature of preparers' justifications, and preparers have been posited to vary or stylize the nature of their workpaper documentation as part of their reputation management strategies (Rich et al. 1997; Gibbins and Trotman 2002).1

Current trends toward real-time reviews involve greater and earlier interactions among reviewers and preparers, where reviewers provide front-end guidance to preparers in the conduct of the audit (Rich et al. 1997; Gibbins and Trotman 2002). One concern arising from real-time reviews is that reviewers communicate their task preferences (that is, their preferred audit conclusions on the task) to workpaper preparers too early in the audit process (Rich et al. 1997; Wilks 2002). Prior studies suggest a cost to real-time reviews in that preparers required to justify to a reviewer with a particular directional preference are systematically influenced by knowledge of their reviewer's preference in terms of their hypothesis generation (Peecher 1996) and evidence interpretation (Wilks 2002). Note that preparers in these studies do not form initial audit judgments or conclusions related to the task before knowing the reviewer's preference.

In a departure from these prior studies, we investigate conditions in which preparers do explicitly state their task preferences (that is, their preferred audit conclusions on the task) before knowing the reviewers' task preferences. Specifically, our first factor—justification requirement—refers to an accountability-inducing requirement (Tetlock 1985a, 1992, 2002) where the preparer is required to justify his/her audit conclusion to a reviewer whose task preference is either similar to or dissimilar from the explicitly stated initial task preference of the preparer. Our motivation for looking at this factor can be viewed from two perspectives. First, an inference from prior research is that the costs associated with real-time review arise in part because preparers now have fewer opportunities to form opinions that are dissimilar from that made by the reviewer (in that preparers form a task preference only after knowing the reviewer's task preference). From this perspective, it is important to know the effects when preparers are required to explicitly form their task preferences prior to the reviewer's task preference being communicated to them. Such a requirement naturally increases the likelihood that the preparers form task preferences that are dissimilar from that formed by the reviewer, and findings on the efficacy of this requirement would be useful to CPA firms considering firm policies or interventions that mitigate limitations of real-time reviews. Second, in practice and particularly during periods when reviewers face the practical constraints of tight reporting deadlines and involvement

---

1 Koonce et al. (1995) find no justification effects on audit procedures and processes (e.g., problem identification, information request) adopted by preparers.
in several concurrent audit assignments, reviewers likely communicate their conclusions or opinions only after preparers have collected evidence and documented their initial conclusions. Thus, there are circumstances when preparers naturally have to form their task preferences before knowing the reviewers’ task preferences. For these reasons, we expand on the dimension of justification requirement, and consider the match or mismatch between the initial task preference of the reviewer and that explicitly stated by the preparer; that is, whether the justification requirement is to a reviewer with a task preference that is similar to or dissimilar from that of the preparer. Because preparers’ responses to these justification requirements are likely constrained by the skill sets they possess, we investigate the moderating effects of two personal attributes: technical knowledge and tacit managerial knowledge (specifically, knowledge on managing their superior’s impressions).

We conducted an experiment in which audit preparers performed an ambiguous audit task with no one single appropriate solution (doubtful accounts task). In the experiment, preparers assessed the appropriateness of the client’s accounting treatment for the allowance for doubtful accounts, and explicitly stated their initial task preferences. They were then informed of their superior’s task preferences, and wrote memos to their superior justifying their final decisions.

Results show that preparers attempt to justify their conclusions in one of three ways: (1) presentation of a greater number of arguments in support of versus against their documented conclusion (net persuasive evidence); (2) consideration of a broad spectrum of perspectives (e.g., client-risk, internal control risk, GAAP, materiality, etc) (breadth of issues); and (3) using language to frame evidence so as to highlight evidence consistent with the documented conclusion and downplay evidence inconsistent with the documented conclusion (evidence framing). We find that preparers are strategic in their use of these different forms of justifications, depending on the context and the personal attributes of the preparers. For the first two justification forms, technical and tacit managerial knowledge moderate the effects of justification requirements. Preparers document more (less) net persuasive evidence when they are required to justify to a reviewer with dissimilar (similar) task preferences, but only when preparers possess high tacit managerial and high technical knowledge. We find this same pattern of results with respect to the documentation of breadth of issues; in addition, we find that preparers with high tacit managerial but low technical knowledge also document greater breadth of issues when they are required to justify to a reviewer with similar task preferences. The first two forms (net persuasive evidence and breadth of issues) involve the listing of relevant evidence related to the auditing issue at hand. In contrast, the third form, evidence framing, relates to the use of language to frame pro and con evidence. We find that the determinants of and directional effects on evidence framing are different from the other two justification forms. In contrast to the other two justification forms, evidence framing is used to a greater extent only when high tacit managerial knowledge preparers are required to justify to a reviewer with similar (as opposed to dissimilar) task preferences. Further analysis suggests that these preparers appear to economize on the enumeration of pro versus con reasons by treating evidence framing as a substitute for net persuasive evidence when justifying to a similar-preference reviewer. We also provide evidence that all three forms of justifications have direct and positive effects on evaluators’ assessments of work quality.

---

2 These observations were gathered from interviews with audit partners and managers from the Big 4 firms. They also noted that depending on the audit task and individual auditor, there would be variations in terms of when the reviewer’s conclusion is communicated to the preparer (i.e., before or after the preparer has formed an initial conclusion).
Our study contributes to the accounting and psychology literature in a number of ways. Prior research in accounting (e.g., Peecher 1996; Wilks 2002) focuses on the construct, justification to a superior with known preferences, and shows that it systematically affects preparers’ cognition or judgment. We expand on this construct, and demonstrate that what matters to the preparer is not simply the reviewer’s task preference per se, but the match or mismatch between the reviewer’s task preference and the initial task preference explicitly stated by the preparer. A similar point can be made to the related prior psychology literature on bolstering and escalating commitment. This literature documents that when people have explicitly stated their initial decision, they have a heightened tendency to escalate commitment to and justify the correctness of that initial decision, even in the face of negative feedback (Staw 1976; Fox and Staw 1979; Caldwell and O’Reilly 1982; Brockner 1992). Our results suggest that the extent of bolstering or escalating commitment to an explicitly stated initial decision may depend on the match or mismatch between this explicitly stated initial decision and the preferences of the justifier.3

Prior research also suggests that knowledge of the reviewers’ task preferences leads to dysfunctional effects on preparers’ performance (Peecher 1996; Wilks 2002). Our study suggests that dysfunctional effects may not necessarily be the case. We provide evidence that dysfunctional effects may occur if the reviewers’ communicated task preferences are similar to that of the preparers, but not when these preferences are dissimilar. In fact, our results suggest that benefits may actually accrue in the latter case with respect to listing “hard” evidence related to pro versus con reasons and auditing/accounting issues.

In addition, while the separate effects on auditors’ performance of technical knowledge and tacit managerial knowledge have been examined (Tan and Libby 1997), there is no empirical evidence on the interactive effects of these two variables, nor how they jointly interact with justification requirements.4 The closest study we are aware of is by Tan and Kao (1999), who show that auditors’ accountability to a superior with unknown preferences interacts with technical knowledge. They do not examine the effect of tacit managerial knowledge. Our results imply that the effects of accountability and technical knowledge are more complex than what has been documented in the extant literature, and that these findings may be relatively incomplete without considering the moderating role of tacit managerial knowledge.

Finally, relatively little research has been done on the determinants of justification memo quality. A recent tax study finds that various aspects of tax knowledge are associated with adaptive information search, which in turn is associated with the quality of justification memos (Magro 2005). However, her study does not examine the specific forms of justifications that may be adopted by tax professionals in documenting research memos, nor does it examine the interactive effects of technical knowledge with tacit managerial knowledge and justification requirements. Our study extends the literature on justifications by documenting the kinds of justifications that preparers naturally engage in, and by providing evidence that the nature and extent of these justifications is a complex function of the attributes of the preparers and justification requirements. We also provide evidence that justification memos using different justification strategies influence memo quality, a finding that has important implications on auditors’ defense of their work quality.

---

3 Much of the escalating commitment literature invariably focuses on dependent variables such as resource allocation assessments (see review by Brockner 1992), rather than justification strategies or justification quality related to technical issues.
4 Tan and Libby (1997) focus on how auditors of different performance ranks (e.g., highly ranked manager versus mediocre manager) have different levels of technical knowledge or tacit managerial knowledge. They do not examine how these two variables interactively determine performance.
The rest of the paper is organized as follows. We develop our hypotheses in Section II, describe the research method in Section III, report the results in Section IV, and conclude with a discussion of the implications and limitations of the paper in Section V.

II. BACKGROUND AND HYPOTHESES DEVELOPMENT

Justification Forms

In formulating a justification, the justifier attempts to provide support for an expressed viewpoint. Put another way, the justifier wants to persuade the target audience that his/her viewpoint is valid. While justifications in auditing settings have been previously investigated (Koonce et al. 1995; Tan et al. 1997), there is little empirical evidence on the forms of justifications employed by workpaper preparers, or features that make a justification persuasive in the eyes of the reviewer.

Justifications can be viewed as a form of reasoned argument, such that a justification is more persuasive the more “reasoned” or logical its argument quality (Petty et al. 1983; Petty and Wegener 1991). For instance, following the jurisprudence model by Toulmin and colleagues (Toulmin 1958; Toulmin et al. 1984), a reasoned argument consists of a claim, along with (a) grounds, considerations, and backings; and (b) rebuttals that acknowledge the presence of certain facts/evidence that may invalidate or weaken the claim.5 Toulmin et al.’s (1984) work suggests that an argument or justification is persuasive if the logical components identified above are present. Conditions (a) and (b) together suggest that listing both reasons supporting (pro) and refuting (con) the claim can be viewed as providing reasoned argumentation. Indeed, the persuasion literature demonstrates that two-sided arguments (involving both pro and con arguments) are more persuasive than one-sided arguments (incorporating only pro arguments) in that the former approach reduces the likelihood of counter-argumentation and enhances message credibility (Pechmann 1992). However, if too many con arguments are included, then the persuasiveness of the message may be eroded (Toulmin et al. 1984; Golden and Alpert 1987). Hence, we expect reasoned arguments to have more pro than con evidence (net persuasive evidence).

Justifications that carefully assess all considerations (pro and con reasons) related to the claim (as suggested by Toulmin et al. [1984]) also likely involve an assessment of a broad variety of issues from different angles. In an auditing context, auditors are trained to arrive at their conclusions after considering the implications of a variety of issues such as the client’s business risk, industry risk, businesses processes and controls, materiality, and GAAP (e.g., see Winograd et al. 2000). All things considered, the basis for the conclusion can be seen to be comprehensive and persuasive the more such issues are considered; we refer to this as breadth of issues.

Other than the listing of evidence (e.g., net persuasive evidence and breadth of issues), argument presentation by way of language influences its persuasiveness (Hirschberg 1990). More recently, Areni (2003) draws upon the literature in logic, social psychology, jurisprudence, and sociolinguistics to suggest that persuasive arguments may be viewed as those that have certain grammatical features in addition to structural features identified by Toulmin et al. (1984). Specifically, justifications may be framed by using language to downplay evidence that is inconsistent with one’s conclusion by countering it with related pieces of consistent evidence (Kamins and Assael 1987; Hirschberg 1990; Crowley and Hoyer 1994; see also Rich et al. 1997). This form of framing has been found to influence auditors’

5 According to Toulmin et al. (1984), claims can have associated qualifiers (e.g., use of words like possibly, certainly, etc.) to reflect the strength of the claim. For example, in the case of some accounting issues such as contingent liabilities, the use of probability phrases is prescribed by authoritative guidance.

The Accounting Review, March 2006
judgments (Tan and Yip-Ow 2001). For example, consider the following sentence justifying the need for additional allowance for doubtful accounts: "notwithstanding the use of experienced credit agents, companies with lower credit ratings are more likely to go bust." A few features of this justification should be noted. First, there is strategic placement of the evidence. The merit of the evidence related to the use of experienced credit agents (inconsistent with the documented conclusion) is downplayed by juxtaposing it with a piece of evidence that is consistent with the desired conclusion (greater allowance for doubtful accounts because accounts receivable may become uncollectible). Second, there is the use of semantics in that the word "notwithstanding" serves to link the inconsistent and consistent evidence, and to downplay the merits of the inconsistent evidence.

Overall, we examine three forms of justifications that may be used by preparers to enhance the persuasiveness of their justifications: net persuasive evidence, breadth of issues, and evidence framing. We discuss determinants of their usage below.

**Determinants of Justification Form Usage**

An extensive body of literature shows that accountability, the need to justify decisions to an audience, generally leads to greater effort and vigilance (Tetlock 1985a, 1992, 2002). Prior accounting studies show that accountability to a superior with unknown preferences has debiasing properties and leads to improved performance (Kennedy 1993; Tan and Kao 1999). Research also shows that the need to justify decisions to a superior with known preferences leads auditors to align their judgments and interpret evidence in a manner consistent with the superior's preferences (Peccher 1996; Turner 2001; Wilks 2002). However, as we posit below, the effects of justification are likely dependent on the match or mismatch between the preparer's explicitly stated initial preference and the preference of the reviewer.

The literature on escalating commitment documents a heightened tendency for individuals to justify the correctness of their explicitly stated initial decisions in the face of negative feedback, particularly when they feel personally responsible for the initial decision (Staw 1976; Fox and Staw 1979; Caldwell and O'Reilly 1982; Brockner 1992). Social psychology research also shows that individuals are more (less) likely to engage in deep cognitive processing when they are exposed to information that is inconsistent (as opposed to consistent) with their initial beliefs (Pyszczynski and Greenberg 1981; Wong and Weiner 1981; Hastie 1984). Individuals exposed to information that is inconsistent with their initial beliefs engage in greater cognitive processing of information to either "explain away" the information or to change their beliefs (Pyszczynski and Greenberg 1981; Wong and Weiner 1981; Hastie 1984). This cognitive effect is likely stronger when the challenge to one's initial belief comes from a superior or evaluator (e.g., a workpaper preparer's reviewer; Tetlock 1985a, 1985b). Moreover, in the auditing context, because the preparer's justification requirement is to a superior who evaluates his/her documented work, we anticipate greater effort to be exerted in the preparation of the justification memo.

---

6 Rich et al. (1997) also identify other strategies such as selective failure to document evidence and evidence fabrication. These strategies involve some potentially unethical behavior of preparers, and fall outside the framework here (based on reasoned argument and evidence framing). Rich et al. (1997) also identify evidence order as a persuasion strategy. However, because it is not clear whether any particular sequence (all pro then con reasons, or vice versa; interweaving pro and con reasons) is deemed to be more persuasive, we do not posit evidence order as a specific justification strategy. We content-analyzed the justifications from participating auditors in our study, and find no evidence that participants order the pro and con reasons in any systematic way.
The direction and manner with which this increased effort is exerted toward documenting persuasive justifications is likely to vary by the auditor's tacit managerial knowledge (henceforth, TMK), which refers to job knowledge related to managing self, others, and career (Wagner and Sternberg 1985). TMK has been associated with superior performance evaluations within the CPA firm (Tan and Libby 1997). The aspect of TMK most relevant here relates to knowledge about managing others, specifically, with respect to managing a superior's impressions of one's work quality. Preparers with greater TMK are likely to be more cognizant of the importance of managing a superior's impressions, and to be more aware of the different justification forms and their efficacies in persuading the reviewer that the documented conclusion and associated justifications are carefully thought out.

In particular, preparers with greater TMK are likely more aware of the greater need to use even more persuasive justifications (through using the justification forms described above) when the reviewer's preferences are dissimilar than when the preferences are similar. This is because, in the former instance, if they choose to remain with their explicitly stated initial position despite the opposing view held by the reviewer, these preparers are likely to anticipate that the reviewer would want to know why their position is more justifiable than that of the reviewer. Even if the preparers were to adopt the position favored by the reviewer, high-TMK preparers are likely more cognizant of the need to invest more effort to employ the different justification forms to elaborate on why they changed from their initially explicitly stated task preferences—this is to impress upon the reviewer that they changed opinions through a reasoned process, and not to merely comply with the reviewer's task preferences. On the other hand, in the case where the preparer-reviewer task preferences are similar, preparers may feel a smaller need to do so because their initial (and final) conclusion is favored by the reviewer.

However, we expect technical knowledge (henceforth, TK) to constrain the extent to which these high-TMK preparers are able to successfully employ more of these three justification forms (net persuasive evidence, breadth of issues, and evidence framing) when the preparer-reviewer preferences are dissimilar. Technical knowledge is needed to employ each of these justification forms (Libby and Luft 1993). Compared to preparers with higher TK, those with lower levels of TK may not be able to successfully use these justification forms even if they wanted to because they may not have the requisite TK to generate the pro and con reasons (for net persuasive evidence and evidence framing), and auditing- and client-related issues (breadth of issues).

---

7 The argument here is that low TMK auditors have less but not no knowledge about these issues.

8 Audit preparers may choose to remain with their initially expressed preference because they are convinced that it is the appropriate conclusion. They may also do so as an impression management technique in that superiors may value subordinates who are capable of defending a chosen outcome more than those who merely agree with what the reviewer prefers. In our study, we asked a Panel of 12 audit partners to assess the importance of various forms of actions and behavior on subordinates’ career advancement prospects. They rated the ability of the subordinates to promote their ideas and convince superiors on the worth of their work (mean rating = 7.58, on a scale of 1 = extremely unimportant to 9 = extremely important) as more important than preparers either: (a) aligning with superior’s convictions of how things need to be done even if it contradicts their own convictions (mean = 5.08, p = 0.009); (b) adjusting workpapers to suit reviewer styles and preferences (mean rating = 5.42, p = 0.001); and (c) being receptive to superior’s preferences (mean = 6.08, p = 0.010, all two-tailed tests).

9 Preparers who switch task preferences to be aligned with their reviewer’s preferences may also be less likely to exert less effort to justify their documented conclusion (in that they know that their final task preferences are favored by the reviewer). However, this proclivity is likely mitigated by their incentive to manage their reputation by demonstrating that they switched preferences for good reasons (and not because they merely followed what the reviewer suggested).
In summary, we predict that preparers with dissimilar preparer-reviewer preferences are more likely to use the three justification forms (net persuasive evidence, breadth of issues, and evidence framing), and that this effect is largest for those with high TMK and high TK. This leads us to the following hypotheses.

**H1a:** Preparers will document more net persuasive evidence (pro less con reasons) to justify their documented conclusions when their initial preferred conclusions are dissimilar from (versus similar to) those of the reviewers, with the difference being higher for preparers with higher levels of TMK and TK.

**H1b:** Preparers will document more breadth of issues to justify their documented conclusions when their initial preferred conclusions are dissimilar from (versus similar to) those of the reviewers, with the difference being higher for preparers with higher levels of TMK and TK.

**H1c:** Preparers will engage in more evidence framing to justify their documented conclusions when their initial preferred conclusions are dissimilar from (versus similar to) those of the reviewers, with the difference being higher for preparers with higher levels of TMK and TK.

### III. RESEARCH METHOD

**Participants**

One hundred fifteen auditors from two Big 4 firms in Singapore participated in this study. They comprised 60 first-year audit staff and 55 third-year audit seniors.

**Independent Variables**

The study employs three independent variables: justification requirement (henceforth, JR) (requirement to justify their conclusion to a reviewer with similar or dissimilar task preferences), technical knowledge (TK), and tacit managerial knowledge (TMK). The first variable, JR, is a between-subjects variable manipulated at two levels. Participants are required to justify their audit conclusion (formed after knowing the reviewer’s task preference) to a reviewer whose initial task preference (i.e., preferred conclusion) is either similar to or dissimilar from that of the participants (that is, their explicitly stated initial conclusion before knowing the reviewer’s preference).

The second independent variable, TK, is a measured variable. The TK test contains five multiple choice questions, and an open-ended question requiring participants to briefly explain how an auditor would assess the adequacy of specific allowance for doubtful accounts. One of the multiple-choice questions pertains to factors to consider in evaluating accounting estimates, another pertains to analytical procedures related to the accounts receivable balance, while the remaining three questions pertain to accounting issues related to the sales and collections cycle of manufacturing clients. The questions are adapted from past CPA exams. With respect to the multiple-choice questions, TK is measured in terms of the number of correct responses obtained. The open-ended question is given the same weight as each of the multiple-choice questions. The response is considered correct if the participant provides a reasonable explanation for assessing the adequacy of allowance for doubtful accounts, and incorrect otherwise.

The third variable, TMK is measured using two-work related audit scenarios. These are developed by reference to Wagner and Sternberg (1985) and Tan and Libby (1997).
Eight options are presented under each scenario, and the rating for each option is on a nine-point scale (1 = extremely unimportant to 9 = extremely important). For the first scenario relating to managing the superior’s impressions, participants are required to rate the importance of learning the preferences and styles of the superior, and being receptive to the superior’s work styles and documentation preferences. For the second scenario relating to ways to progress up the firm’s hierarchy, participants are required to rate the importance of strategies such as providing feedback to the superior on one’s work progress, seeking feedback from the superior, socializing with the superior in non-work-related activities, drawing awareness to one’s achievements, and promoting one’s ideas and work.

Twelve partners (mean experience = 18 years) from two of the participating firms completed the TMK instrument. The partners’ responses serve as the benchmark against which the preparers’ TMK is measured. TMK is measured as the sum of the squared deviations of a participant’s ratings from the mean ratings of the Panel of partners. Subsequently, the squared deviation score of each participant is subtracted from the participant who obtained the highest score (that is, the worst score, 171.97), such that the higher the score, the better the participant’s TMK.

Task

The doubtful accounts task is adapted from Hackenbrack and Nelson (1996) and Tan and Jamal (2001). It is set in the context of Innovate, a company that has been the client of the audit firm for the past five years. The company specializes in producing sound cards for personal computers. During the fourth quarter of the past year, the company adopts a more liberal credit policy in an attempt to meet its sales target. In order to manage the risk from its new credit policy, the company hires two experienced “credit agents” to monitor the collectability of the accounts. Hence, the financial controller feels that no additional allowance for doubtful accounts is needed. However, he is willing to provide a detailed footnote on the above matter. Currently, the company is in need of additional financing to meet its operational needs, and the bank requests an audited set of financial statements to discuss the terms of the loan. The case materials include the profit and loss account with details on sales, cost of sales, gross margin, bad debt expenses and other expenses for the current and prior years, and the debtors’ aging analysis for the current year. Given this background, participants are to make an assessment on whether to permit a footnote disclosure or to ask for additional allowance for doubtful accounts.

Procedure

The experiment consisted of two stages, and was administered during in-house training sessions conducted by the participating firms, in the presence of a training manager. To enable us to track the participants across the two stages, and to induce accountability and motivation, we requested the participants to write down their names and contact numbers on the research instrument. In Stage 1, the participants were presented with two sets of material placed in separate sealed envelopes to ensure that the participants performed the task in the intended sequence. The first envelope contained the task materials. Participants first read the case details, assessed their agreement with the client’s accounting treatment

---

10 To ensure that the items selected for inclusion in the tacit managerial knowledge measure have substantial agreement among the partners with regard to their level of importance, the partners’ responses to the 16 items are first evaluated, and only items that have a range of 5 or less are selected. This results in the selection of ten items. These items have, on average, lower standard deviation (1.19) and skewness (−0.80), than the remaining 6 items (s.d. = 2.07, skewness = −1.39). This approach is similar to that adopted in Tan and Libby (1997).

The Accounting Review, March 2006
on a scale of 1 (definitely insist on additional allowance) to 11 (definitely agree with footnote disclosure), made a decision on whether to insist on additional allowance or permit a footnote disclosure, and then indicated their confidence in the decision. After that, the participants proceeded to the second envelope containing some problem-solving ability questions (based on Tan and Kao [1999]) and background information questions.

Stage 2 of the experiment was conducted on a subsequent day of the participants’ training. In Stage 2, the participants were presented with two sets of materials placed in separate sealed envelopes. The cover sheet on the first envelope informed the participants that an audit manager had agreed to participate in the exercise by acting as the audit manager in charge of the audit of Innovate. They were then informed of the need to write a memo to the manager justifying their decision on the accounting treatment to be used by the client. Finally, the instructions stated that, as in an audit review, an audit manager would review the memos. The first envelope contained the task with the reviewer’s task preferences (i.e., whether the reviewer agreed or disagreed with the client’s accounting treatment) indicated on a separate sheet. The instructions stated that the reviewer had briefly reviewed the case details and made a preliminary judgment. The participants first read the reviewer’s task preferences, re-assessed the appropriateness of the client’s accounting treatment with the knowledge of the reviewer’s task preferences (on the same scale as that used in Stage 1), and indicated their decisions on the appropriateness of the client’s accounting treatment together with their confidence in the decision. Subsequently, they wrote a memo to the reviewer on a separate sheet justifying their decision. The case materials (without the participants’ responses to Stage 1) were included in the envelope for their reference. Finally, participants proceeded to another envelope that contained the manipulation check, TMK, and TK questions.

Dependent Variables

The dependent variables for H1a, H1b, and H1c are the specific forms of preparers’ justifications. The first form is net persuasive evidence (H1a), which we define as the number of pro minus con reasons related to a documented conclusion. The second form, breadth of issues (H1b), relates to the issues covered in the preparers’ justifications, and is measured by computing the sum of different auditing and accounting issues (e.g., internal control assessment, client related factors, financial statement analysis, industry-related factors, GAAP, and materiality) mentioned in the preparers’ justification memos. We elicited these issues from an audit manager and two colleagues with auditing experience during a pre-test. The third form is evidence framing (H1c), which we identify by observing whether evidence is framed toward the preparers’ conclusions (that is, by using semantics to highlight evidence consistent with their conclusions and downplay evidence inconsistent with their conclusions). The Appendix provides examples of the various forms of justification.

---

11 The participants also performed an additional ambiguous audit task of higher complexity related to hedging of foreign currency. Task order was counterbalanced, and order does not moderate our findings (p > 0.340). For this higher complexity task, participants had to assess whether it was appropriate for a holding company to hedge on behalf of its wholly owned subsidiary and, hence, reflect the exchange differences arising on the foreign currency liability in its foreign exchange reserves account. Participants found this task very complex, and wrote very few justifications. For example, only four participants engaged in evidence framing, and, in the dissimilar preparer-reviewer preference condition, only ten participants wrote any form of justification. This resulted in cell sizes that are too small for meaningful statistical tests on the effects of JR, TK, and TMK.

12 Inferences relating to the hypotheses are unchanged with the inclusion of problem-solving ability as a control variable.

13 Among those who used evidence framing, other than one participant, all engaged in only a single instance of evidence framing. Hence, we coded evidence framing as 1 if there was at least one instance of evidence framing, and 0 otherwise.
Determinants of Audit Preparers' Workpaper Justifications

One of the authors and an independent coder with three years of auditing experience content-coded the justifications documented in the participants' memos. The coders were blind to the experimental treatments (that is, the reviewer's task preferences). The attributes coded are the number of pro and con reasons related to a documented conclusion, the audit and accounting issues identified (e.g., internal control assessment, client-related factors, financial statement analysis, industry-related factors, GAAP, and materiality), and the use of evidence framing. In coding the justification forms, disagreements between the two coders were resolved through discussion. The inter-rater agreement (Kappa coefficient) is high: 0.886 (pro and con statements), 0.930 (breadth of issues), and 0.862 (evidence framing, \( p = 0.000 \) in all instances).

IV. RESULTS

Preliminary Analyses and Manipulation Checks

As a manipulation check, we ask participants whether the judgment of the audit manager agreed or disagreed with the client. The majority correctly answered this question, and the responses of ten participants who failed the manipulation check are excluded from the analyses. Hence, the final sample consists of 105 usable responses. We also ask participants to indicate the extent to which they expect the memos to be reviewed by an audit manager (1 = strongly disagree; 9 = strongly agree). The mean rating of 7.33 suggests that, on average, participants expect the memos to be subsequently reviewed by an audit manager.

The mean TK score is 3.66 (varying from 1 to 6), and the mean TMK score is 139.83 (varying from 0 to 166.36). Cronbach's alpha for the TMK questions is 0.79. The two independent variables are not significantly correlated (Pearson \( r = -0.115, p = 0.245 \), two-tailed test).

To obtain a proxy for effort, we asked participants to indicate the time they started and ended the task. Our theory predicts that those who are in the dissimilar-preference condition are more likely to exert greater effort than those in the similar-preference condition. Consistent with our expectations, we find that the dissimilar-preference group spent more time writing the memo (mean = 8.74 minutes) than the similar-preference group (mean = 7.70 minutes, \( p = 0.088 \), one-tailed test).

We aggregate the analyses across firms as the results are not moderated by firm type \( (p > 0.606) \).

Descriptive Statistics on Forms of Justification

Panels A and B in Table 1 present the descriptive statistics on the different justification forms that preparers use to justify their documented conclusion to a reviewer with either similar or dissimilar task preferences. The net persuasive evidence (pro less con arguments) ranges from a low of -1 to a high of 6, with an average of 1.63 (1.77) for preparers required to justify to a reviewer with similar (dissimilar) task preferences. The breadth of issues ranges from a low of 0 to a high of 6 issues, with an average of 2.17 (2.15) issues.

---

14 Inferences relating to the hypotheses are unchanged with the inclusion of the responses of participants who failed the manipulation check.
15 Due to the multidimensional nature of the technical knowledge construct, we do not report any measure of reliability (see Bonner and Walker 1994).
16 48 (45.7 percent) of the participants favor the sufficiency of footnote disclosure, while 57 participants (54.3 percent) favor the need for additional allowance in their initial decisions. Decision type (footnote versus allowance) does not moderate our findings \( (p > 0.229) \).
TABLE 1
Descriptive Statistics on Forms of Justification
(Net Persuasive Evidence, Breadth of Issues, Evidence Framing)

<table>
<thead>
<tr>
<th>Panel A: Forms of Justification Used by Preparers Required to Justify Their Documented Conclusion to a Reviewer with Similar Task Preferences (n = 52)</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net persuasive evidence</td>
<td>-1</td>
<td>6.00</td>
<td>1.63</td>
<td>1.21</td>
</tr>
<tr>
<td>Breadth of issues</td>
<td>0</td>
<td>6.00</td>
<td>2.17</td>
<td>0.98</td>
</tr>
<tr>
<td>Evidence framing</td>
<td>0</td>
<td>1.00</td>
<td>0.44</td>
<td>0.50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Forms of Justification Used by Preparers Required to Justify Their Documented Conclusion to a Reviewer with Dissimilar Task Preferences (n = 53)</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net persuasive evidence</td>
<td>-1</td>
<td>4.00</td>
<td>1.77</td>
<td>1.17</td>
</tr>
<tr>
<td>Breadth of issues</td>
<td>0</td>
<td>6.00</td>
<td>2.15</td>
<td>1.03</td>
</tr>
<tr>
<td>Evidence framing</td>
<td>0</td>
<td>1.00</td>
<td>0.40</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Net persuasive evidence refers to the number of pro arguments less the number of con arguments related to a documented conclusion. Breadth of issues is the sum of accounting and auditing issues the participants consider in justifying their conclusions. Evidence framing refers to whether participants frame their evidence to highlight evidence consistent with the documented conclusion and downplay evidence inconsistent with the documented conclusion.

for preparers required to justify to a reviewer with similar (dissimilar) task preferences.\(^{17}\)

Finally, on average, 44 (40) percent of the participants engage in evidence framing when they are required to justify to a reviewer with similar (dissimilar) task preferences.

Test of Hypotheses

To test H1a, H1b, and H1c, separate regression analyses are run for each of the three justification forms, with JR, TMK, TK, as well as their interaction terms as independent variables. JR is coded as 0 when preparer-reviewer initial task preferences are similar, and 1 when dissimilar. Both TMK and TK are centered on their means to avoid multicollinearity with their interaction terms (Aiken and West 1991). Results of these regression analyses are shown in Table 2, Panels A to C.

Hypotheses 1a, 1b, and 1c predict that preparers will use more net persuasive evidence, breadth of issues, and evidence framing, respectively, when their initial preferred conclusions are dissimilar from (versus similar to) those of the reviewers, with the difference being higher for preparers with higher levels of TMK and TK. Table 2, Panel A, shows that the JR × TMK × TK term for net persuasive evidence is marginally significant (p = 0.094, one-tailed test), providing weak evidence of a three-way interaction.

To assess the nature of the interaction, we follow the procedure suggested in Aiken and West (1991), and test simple effects of JR at high (one standard deviation above mean) and low (one standard deviation below mean) levels of TMK and TK. As seen from Table 2, Panel D, when both TMK and TK are high, there is a significantly positive effect of JR on net persuasive evidence (p = 0.014, one-tailed test). The positively significant slope suggests that preparers document more pro arguments than con arguments when they are high.

\(^{17}\) Participants generally identified issues relating to internal control risk (77.1 percent), trends/fluctuations in the key accounting numbers (56.2 percent), GAAP (23.8 percent), and materiality (22.9 percent).

The Accounting Review, March 2006
TABLE 2
Regression Results of the Effects of Justification Requirement, Preparers’ Technical Knowledge and Tacit Managerial Knowledge on Forms of Workpaper Justifications
(Net Persuasive Evidence, Breadth of Issues, and Evidence Framing)

\( (n = 105) \)

Panel A: Net Persuasive Evidence as Dependent Variable

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>( t )</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.614</td>
<td>0.162</td>
<td>9.971</td>
<td>0.000</td>
</tr>
<tr>
<td>JR</td>
<td>0.182</td>
<td>0.229</td>
<td>0.796</td>
<td>0.428</td>
</tr>
<tr>
<td>TK</td>
<td>-0.011</td>
<td>0.143</td>
<td>-0.079</td>
<td>0.937</td>
</tr>
<tr>
<td>TMK</td>
<td>-0.008</td>
<td>0.005</td>
<td>-1.490</td>
<td>0.139</td>
</tr>
<tr>
<td>JR ( \times ) TK</td>
<td>0.292</td>
<td>0.186</td>
<td>1.569</td>
<td>0.120</td>
</tr>
<tr>
<td>JR ( \times ) TMK</td>
<td>0.009</td>
<td>0.009</td>
<td>0.922</td>
<td>0.359</td>
</tr>
<tr>
<td>TK ( \times ) TMK</td>
<td>0.000</td>
<td>0.004</td>
<td>-0.106</td>
<td>0.916</td>
</tr>
<tr>
<td>JR ( \times ) TK ( \times ) TMK</td>
<td>0.008</td>
<td>0.006</td>
<td>1.329</td>
<td>0.094*</td>
</tr>
</tbody>
</table>

Panel B: Breadth of Issues as Dependent Variable

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>( t )</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.161</td>
<td>0.135</td>
<td>16.034</td>
<td>0.000</td>
</tr>
<tr>
<td>JR</td>
<td>-0.003</td>
<td>0.191</td>
<td>-0.014</td>
<td>0.989</td>
</tr>
<tr>
<td>TK</td>
<td>-0.265</td>
<td>0.119</td>
<td>-2.226</td>
<td>0.028</td>
</tr>
<tr>
<td>TMK</td>
<td>0.003</td>
<td>0.005</td>
<td>0.669</td>
<td>0.505</td>
</tr>
<tr>
<td>JR ( \times ) TK</td>
<td>0.532</td>
<td>0.155</td>
<td>3.426</td>
<td>0.001</td>
</tr>
<tr>
<td>JR ( \times ) TMK</td>
<td>0.000</td>
<td>0.008</td>
<td>-0.011</td>
<td>0.991</td>
</tr>
<tr>
<td>TK ( \times ) TMK</td>
<td>-0.004</td>
<td>0.003</td>
<td>-1.220</td>
<td>0.225</td>
</tr>
<tr>
<td>JR ( \times ) TK ( \times ) TMK</td>
<td>0.008</td>
<td>0.005</td>
<td>1.766</td>
<td>0.041*</td>
</tr>
</tbody>
</table>

Panel C: Evidence Framing as Dependent Variable

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Wald</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.367</td>
<td>0.332</td>
<td>1.225</td>
<td>0.268</td>
</tr>
<tr>
<td>JR</td>
<td>-0.072</td>
<td>0.439</td>
<td>0.027</td>
<td>0.870</td>
</tr>
<tr>
<td>TK</td>
<td>-0.095</td>
<td>0.287</td>
<td>0.110</td>
<td>0.740</td>
</tr>
<tr>
<td>TMK</td>
<td>0.044</td>
<td>0.018</td>
<td>5.782</td>
<td>0.016</td>
</tr>
<tr>
<td>JR ( \times ) TK</td>
<td>0.127</td>
<td>0.356</td>
<td>0.126</td>
<td>0.722</td>
</tr>
<tr>
<td>JR ( \times ) TMK</td>
<td>-0.043</td>
<td>0.023</td>
<td>3.420</td>
<td>0.064</td>
</tr>
<tr>
<td>TK ( \times ) TMK</td>
<td>-0.012</td>
<td>0.013</td>
<td>0.912</td>
<td>0.340</td>
</tr>
<tr>
<td>JR ( \times ) TK ( \times ) TMK</td>
<td>0.009</td>
<td>0.015</td>
<td>0.369</td>
<td>0.272*</td>
</tr>
</tbody>
</table>

(continued on next page)

The Accounting Review, March 2006
TABLE 2 (Continued)

Panel D: Tests of Simple Slopes of Justification Requirement (JR) at Different Levels of Preparers' Technical Knowledge (TK) and Tacit Managerial Knowledge (TMK)

|                          | Net Persuasive Evidence as Dependent Variable |  |  |  |  | Breadth of Issues as Dependent Variable |
|--------------------------|----------------------------------------------|---|---|---|---|
|                          | Magnitude of Simple Slope | Std. Error | t   | p-value | Magnitude of Simple Slope | Std. Error | t   | p-value |
| High TK and High TMK     | 1.059                         | 0.473       | 2.237 | 0.014*   | 0.961                         | 0.394       | 2.438 | 0.009*   |
| High TK and Low TMK      | 0.043                         | 0.422       | 0.103 | 0.918     | 0.376                         | 0.352       | 1.069 | 0.288     |
| Low TK and High TMK      | -0.212                        | 0.465       | -0.456 | 0.650     | -0.970                        | 0.387       | -2.507 | 0.014     |
| Low TK and Low TMK       | -0.161                        | 0.496       | -0.325 | 0.746     | -0.376                        | 0.413       | -0.911 | 0.365     |

* One-tailed test.

JR = requirement to justify to a reviewer with similar or dissimilar task preferences (0 = similar task preferences, 1 = dissimilar task preferences);

TK = technical knowledge (measured as the number of correct responses scored by the participants on the knowledge test; High TK is 1 standard deviation above the mean TK, and Low TK is 1 standard deviation below the mean TK); and

TMK = tacit managerial knowledge (measured as the sum-of-squared deviations of participants’ ratings from the mean ratings of the Panel of partners on the TMK questions. Each participant’s TMK score was then subtracted from the participant with the highest score (171.97, that is the worst score). Hence, the higher the TMK score the better is the participants’ TMK; High TMK is 1 standard deviation above the mean TMK, and Low TMK is 1 standard deviation below the mean TMK).

Both TK and TMK are centered on their means to avoid multicollinearity with their interaction terms.

required to justify to a reviewer with dissimilar task preferences than when they are required to justify to a reviewer with similar task preferences. There is no statistically significant effect of JR when either TK, TMK, or both TK and TMK are low (p > 0.650, two-tailed test). These results are consistent with H1a.18

Table 2, Panel B, shows a significant three-way interaction of JR, TMK, and TK on the breadth of issues (p = 0.041, one-tailed test). As seen from Table 2, Panel D, the results of the tests of simple effects of JR are largely similar to that observed earlier. The results show a statistically significant positive effect of JR when both TMK and TK are high (p = 0.009, one-tailed test), but not when TK is high and TMK low, or when both TK and TMK are low (p > 0.288, two-tailed test). However, contrary to expectation, when preparers possess low TK and high TMK, there is a significantly negative effect of JR on the breadth of issues documented (p = 0.014, two-tailed test). Preparers with low TK and high TMK document more issues when they are required to justify to a reviewer with similar task

18 We separately analyze the pro and con reasons, and find a three-way interaction for pro reasons (p = 0.042). For con reasons, the only statistically significant effect we find is a positive main effect of TMK (p = 0.040).

The Accounting Review, March 2006
preferences than when they are required to justify to a reviewer with dissimilar task preferences.\textsuperscript{19} These results are partially consistent with H1b.\textsuperscript{20}

Table 2, Panel C, presents the results of the logistic regression analysis on the presence or absence of evidence framing. Contrary to expectations, we find no statistically significant three-way interaction effect of JR × TMK × TK (p = 0.272, one-tailed test). Instead, we find a marginally significant two-way JR × TMK interaction (p = 0.064, two-tailed test). Further analysis reveals a marginally significant negative effect of JR when TMK is high (p = 0.069, two-tailed test) but not when TMK is low (p = 0.219, two-tailed test). High-TMK preparers are more likely to use evidence framing when they are required to justify to a reviewer with similar task preferences than when they are required to justify to a reviewer with dissimilar task preferences. TK does not moderate this effect, nor does it interact with any other variable. The absence of a TK effect may be related to the fact that among those who engaged in evidence framing, all but one engaged in only a single instance of evidence framing (i.e., involving only one pair of pro and con reasons, but not more than that). Thus, the level of TK required to generate the single pair of pro and con reason may not be high, and most participants likely have the requisite baseline knowledge.

The directional effect of JR—high-TMK preparers use more evidence framing in a similar task preference JR condition than in a dissimilar task preference JR condition—is also opposite to what we anticipated and found for the other two justification forms. We speculate that one reason is that high-TMK preparers may view evidence framing as an efficient yet effective manner to convince a similar-preference reviewer on the quality of their justifications.\textsuperscript{21} High-TMK preparers may be aware of the need to show a two-sided argument involving both pro and con reasons (Toulmin et al. 1984; Golden and Alpert 1987), but may also be aware that with respect to a similar-preference reviewer, they may be able to economize on the extent to which they need to do so. Therefore, they conserve their effort toward efficiently using language to frame evidence that involves downplaying a single con reason through emphasizing a single pro reason (rather than engage in enumeration of a range of net persuasive evidence and breadth of issues). Consistent with this “efficiency” perspective, recall that literally all preparers using evidence framing engaged only in one (not multiple) instance of evidence framing. Also, within the similar-preference group, the use of evidence framing is negatively associated with documentation of net persuasive evidence (Pearson r = −0.571, p = 0.000, two-tailed test). This is consistent

---

\textsuperscript{19} It could be that these high-TMK preparers perceive that, given their low TK, they would be less able to generate sufficient breadth of issues needed to persuade a reviewer with dissimilar task preferences (who might require substantial breadth of issues to be persuaded). Thus, they focus their effort on a reviewer with similar task preferences. It may also be that TMK compensates for the lack of TK in this situation.

\textsuperscript{20} In coding breadth of issues documented, different aspects of the same issue (e.g., financial statement analysis) are coded as one issue, even though these would have been considered separate pro or con reasons. To assess the sensitivity of our results to this coding, we re-analyze the results using the sum of pro and con reasons as dependent variable, and obtain similar results as those using breadth of issues as dependent variable.

\textsuperscript{21} We find no significant difference in confidence between participants required to justify to a reviewer with similar preferences and those required to justify to a reviewer with dissimilar preferences (means = 75.87 and 75.40, respectively, p = 0.842, two-tailed test). Furthermore, inferences relating to the hypotheses remain unchanged with the inclusion of participants’ confidence in their decision as a control variable.
with the argument that evidence framing is viewed as a substitute for net persuasive evidence. These findings suggest that preparers strategically use different justification forms in reaction to justification requirements, conditional on the skill sets they possess. Overall, our predictions are generally supported for two (net persuasive evidence and breadth of issues) out of the three forms of justifications.

Effects of Maintaining versus Switching One’s Initial Task Preferences

We also perform additional analyses on the 53 (out of 105) participants who were required to justify to a reviewer with dissimilar task preferences. Specifically, we investigate whether there are any differences in the nature and extent of justification forms documented by the participants who changed their decision to align with the reviewer’s expressed task preference (“switch” group; n = 15, 28 percent) and those who remained with their original decision (“no-switch” group; n = 38, 72 percent). As expected, we observe no significant differences between the two groups in terms of the use of net persuasive evidence (means = 1.60 and 1.84 for the switch and no-switch groups, respectively, p = 0.503, two-tailed test) and breadth of issues (means = 2.20 and 2.13 for the switch and no-switch groups, respectively, p = 0.829, two-tailed test). However, we find that preparers who switch task preferences (and end up with final task preferences similar to that by the reviewer) have a marginally greater propensity to use evidence framing than those who do not switch (means = 0.60 and 0.32 for the switch and no-switch groups, respectively, p = 0.062, two-tailed test). This result is somewhat consistent with our finding in our main analyses indicating that evidence framing is more likely used when preparer-reviewer task preferences are similar. It is also possible that evidence framing is used as a means for these preparers to acknowledge “correction” of their initial task preference by highlighting how a pro reason (in favor of their revised decision) overrode initial concerns about a con reason.

Effects of Justification Forms on Assessments of Memo Quality

Does greater documentation of each of the three justification forms lead to better assessments of memo quality? Presumably, the answer is “yes” because this is likely why different preparers engage in differential justification efforts. Auditor respondents in Emby and Gibbins’ (1988) survey do perceive that justifiability of a decision is a characteristic of good judgment.

To address this question, we asked two evaluators to assess the quality of the justification memos written by our participants. The evaluators comprised one audit manager with 11 years of audit experience (including experience in the audit training department), and

---

22 In contrast, within the dissimilar-preference group, there is no statistically significant association between the use of net persuasive evidence and evidence framing (r = -0.208, p = 0.136, two-tailed test). Furthermore, the negative correlation between the above two justification forms (net persuasive evidence and evidence framing) is significantly higher in the similar-preference group than in the dissimilar-preference group (p = 0.05, Fisher r-to-z transformation test). Across all participants, between net persuasive evidence and evidence framing, r = -0.393 (p = 0.000); between net persuasive evidence and breadth of issues, r = 0.479 (p = 0.000); and between evidence framing and breadth of issues, r = 0.249 (p = 0.010, all two-tailed).

23 Within the switch group, there is no statistically significant association between the use of evidence framing and net persuasive evidence (Pearson r = -0.302, p = 0.275, two-tailed test). Furthermore, within the dissimilar-preference group, there is no significant difference in confidence levels between participants who change decisions to align with the reviewer and those who remain with their initial decisions (means = 76.67 and 74.89, respectively, p = 0.655, two-tailed test). Finally, within the switch group, we find no main or interactive effects of TK or TMK (p > 0.450, two-tailed test). These tests may lack power because of the small sample size (n = 15).
one professor who was a former audit manager with ten years of audit experience. Evaluators were required to rate the quality of each of the preparers’ justification memos on a scale of 0 (low quality) to 10 (excellent quality). To ensure that the memo quality is unaffected by the legibility of the preparers’ handwriting, the handwritten justification memos are transcribed before being assessed by the reviewers.

Because the task is an ambiguous audit task for which audit quality is primarily determined by the justifiability and defensibility of the documented conclusion, the evaluators were given explicit instructions to consider the quality of the justifications documented by the preparers in support of their stated conclusion, and not base their evaluations on the appropriateness of the documented conclusion without evaluating the quality of the justifications. Both evaluators had different initial task preferences (one favored footnote disclosure, and the other favored additional allowance). There is no significant difference in the mean ratings awarded by Evaluator A to the preparers’ memos with conclusions that were in agreement (mean = 5.46) and those that were in disagreement with the evaluator’s preference (mean = 5.10, p = 0.198, two-tailed test), suggesting no significant effect of conclusion congruence on these ratings. However, Evaluator B gave significantly higher ratings (p = 0.002, two-tailed test) to preparers’ memos with conclusions that disagreed with that of the evaluator (mean = 5.93) and lower ratings to memos with conclusions that agreed with the evaluator (mean = 4.73); this is directionally opposite to what is expected if the evaluator had merely rewarded conclusion congruence. Thus, the evaluator’s ratings appear to be influenced by the quality of the justifications rather than by conclusion appropriateness or congruence. The scores given by the two evaluators are significantly positively correlated (Pearson r = 0.429, p = 0.000, two-tailed test). Paired t-tests also show no significant differences in the average ratings accorded by the two evaluators to each memo (p = 0.663, two-tailed test). We use the average of both evaluators’ ratings as dependent variable.

Table 3, Panel A, reports the results of the regression of the three justification forms (independent variables) on evaluated performance (dependent variable). As expected, all three forms of justifications are positively associated with evaluated performance (net persuasive evidence, p = 0.015; breadth of issues, p = 0.006; evidence framing, p = 0.001, all one-tailed). To assess whether the results hold after controlling for the effects of JR, TMK, and TK, we re-perform the analyses above with JR, TMK, and TK, along with their interactions, as control variables. As shown in Table 3, Panel B, all three forms of justification are again positively associated with evaluated performance (net persuasive evidence, p = 0.020; breadth of issues, p = 0.008; evidence framing, p = 0.001, all one-tailed).24,25

24 Together, the results satisfy the following two conditions listed by Kenny et al. (1998) to establish that justification form mediates the effects of JR, TK, and TMK on evaluated performance: (a) JR, TK, and TMK affect justification form (which we find in the main analyses; Table 2, Panels A to C), and (b) in turn, the mediator (i.e., justification form) affects evaluated performance (Table 3, Panel B), even after controlling for the presence of JR, TK, and TMK. For condition (b), if the three-way interaction effect (JR × TK × TMK) is statistically insignificant, a total mediation occurs. In our case, the interaction is marginally significant (p = 0.052, two-tailed test), which suggests that justification form partially mediates the effects of JR, TK, and TMK on evaluated performance. We also re-perform the regression analysis with evaluated performance as dependent variable, and JR, TMK, TK, and their interactions as independent variables (i.e., without the three justification forms as independent variables). As in our test of H1a, we find a statistically significant three-way interaction (p = 0.010, two-tailed test), and a statistically significant simple main effect of JR only when both TMK and TK are high (p = 0.015, two-tailed test), but not when either TK, TMK, or both TK and TMK are low (p > 0.207, two-tailed test).

25 The variance inflation factor (VIF) for each of the three justification forms is below 2.2, which is less than the cut-off point of 10 that is indicative of multicollinearity problems (Kutner et al. 2004).

The Accounting Review, March 2006
TABLE 3
Evaluators' Assessment of Preparers' Task Performance (Memo Quality)
(n = 105)

<table>
<thead>
<tr>
<th></th>
<th>Regression Coefficient</th>
<th>Standard Error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.457</td>
<td>0.301</td>
<td>11.479</td>
<td>0.000</td>
</tr>
<tr>
<td>Net persuasive evidence</td>
<td>0.318</td>
<td>0.143</td>
<td>2.220</td>
<td>0.015*</td>
</tr>
<tr>
<td>Breadth of issues</td>
<td>0.415</td>
<td>0.161</td>
<td>2.580</td>
<td>0.006*</td>
</tr>
<tr>
<td>Evidence framing</td>
<td>1.036</td>
<td>0.310</td>
<td>3.339</td>
<td>0.001*</td>
</tr>
</tbody>
</table>

Panel B: Regression Results of the Effects of Forms of Justification on Evaluators' Assessment of Preparers' Task Performance (Memo Quality) (Controlling for the Effects of Justification Requirement, Preparers' Technical Knowledge, and Tacit Managerial Knowledge)

<table>
<thead>
<tr>
<th></th>
<th>Regression Coefficient</th>
<th>Standard Error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.255</td>
<td>0.336</td>
<td>9.694</td>
<td>0.000</td>
</tr>
<tr>
<td>JR</td>
<td>0.461</td>
<td>0.238</td>
<td>1.939</td>
<td>0.055</td>
</tr>
<tr>
<td>TK</td>
<td>-0.072</td>
<td>0.153</td>
<td>-0.470</td>
<td>0.639</td>
</tr>
<tr>
<td>TMK</td>
<td>0.000</td>
<td>0.006</td>
<td>0.029</td>
<td>0.977</td>
</tr>
<tr>
<td>JR × TK</td>
<td>-0.044</td>
<td>0.204</td>
<td>-0.218</td>
<td>0.828</td>
</tr>
<tr>
<td>JR × TK</td>
<td>0.002</td>
<td>0.010</td>
<td>0.178</td>
<td>0.859</td>
</tr>
<tr>
<td>TK × TMK</td>
<td>-0.005</td>
<td>0.004</td>
<td>-1.441</td>
<td>0.153</td>
</tr>
<tr>
<td>JR × TK × TMK</td>
<td>0.012</td>
<td>0.006</td>
<td>1.969</td>
<td>0.052</td>
</tr>
<tr>
<td>Net persuasive evidence</td>
<td>0.304</td>
<td>0.145</td>
<td>2.089</td>
<td>0.020*</td>
</tr>
<tr>
<td>Breadth of issues</td>
<td>0.409</td>
<td>0.166</td>
<td>2.456</td>
<td>0.008*</td>
</tr>
<tr>
<td>Evidence framing</td>
<td>1.058</td>
<td>0.313</td>
<td>3.375</td>
<td>0.001*</td>
</tr>
</tbody>
</table>

* One-tailed test.

Evaluators' task performance (memo quality) is assessed by two evaluators (a manager with a Big 4 accounting firm, and an academic with prior experience as an audit manager), on a scale of 0 (low quality) to 10 (excellent quality). The average of their ratings serves as the dependent variable.

V. CONCLUSION

We identify three ways by which preparers attempt to justify their documented conclusions. We provide the first empirical evidence that the form and extent of justifications vary by whether preparers are required to justify their explicitly stated conclusions to a reviewer with similar or dissimilar initial task preferences, and the preparers’ tacit managerial knowledge and technical knowledge. We find that high tacit managerial knowledge preparers are strategic in terms of the circumstances under which they use particular forms of justifications. They have a greater tendency to use language persuasively in framing their evidence (i.e., evidence framing) when they are required to justify to reviewers with similar initial task preferences. In contrast, preparers with high tacit managerial knowledge and high technical knowledge have a tendency to articulate and enumerate audit evidence (e.g., listing of pro versus con reasons; breadth of issues considered) to provide logical support for the stated conclusion when they are required to justify their conclusions to reviewers with dissimilar initial task preferences. Thus, the nature and extent of justifications used vary

The Accounting Review, March 2006
by environmental factors (here, the preferences of the reviewer) and attributes of the preparer (here, the preferences of the preparer, and his/her tacit managerial knowledge and technical knowledge).

Readers should note some limitations of this study. First, the justification forms that we investigate are likely a subset of the repertoire of justification approaches that preparers use. For example, consultation has been documented to be a means of improving the justifiability of a decision (Kennedy et al. 1997), but this option was not available to our participants because discussion was not permitted in our experiment. Second, our participants are staff auditors and seniors, and their justification approaches may differ from those of managers, who may have more sophisticated ways of dealing with reviewers’ preferences. Third, we consider a setting where the reviewer’s preferences influence the work done by preparers. It may well be that the reverse is true—the preparers’ work influences the reviewer’s preferences, which in turn influences the work performed by preparers. Our study does not capture this possible interactive relation, which is an avenue for future research. Fourth, in evaluating the participants’ memos, evaluators were given explicit instructions to consider only the quality of the justifications. This quality measure we obtain may not generalize to what occurs in practice because without such instructions, evaluators’ ratings may also be influenced by the congruence of the documented conclusion with their preferred conclusion rather than by the justification’s quality. Finally, our current study uses Singapore auditors as participants, and the results may not generalize to U.S. auditors. This concern is mitigated by the fact that our participating auditors are from the Big 4 accounting firms, all of which adopt consistent global practices. Prior studies using Singaporean auditor participants have also documented results that are consistent with those using U.S. auditors (Tan and Libby 1997; Tan and Kao 1999).

We conclude by highlighting some implications of our findings. One implication is that there are potential benefits that may accrue to the audit review process from having preparers form and explicitly state their task preferences before becoming aware of the reviewer’s task preferences, especially when preparer-reviewer task preferences are dissimilar. Furthermore, if the reviewer wants to motivate the preparer into thinking harder about net persuasive evidence and breadth of issues to support the documented conclusion, the reviewer could, early in the audit process, play the role of “devil’s advocate” (Schwenk and Valacich 1994). The reviewer could first find out or anticipate the preparers’ task preferences, and then communicate an initial task preference that is opposite to that held by the preparer. Similarly, the reviewer could invoke a counter-argument mind-set in preparers by prodding them to consider potential alternatives (Kray and Galinsky 2003). However, for these mechanisms to be effective, preparers need to possess the requisite tacit managerial knowledge and technical knowledge.

Our finding that the panel of evaluators (comprising a senior audit manager and a former audit manager) assessed justification memos using these three justification forms to be of higher quality has practical importance. Extant literature argues that decision justifiability is a characteristic of good audit judgment (Emby and Gibbins 1988), but the relation between justifications and assessed work quality has hitherto not been established. To the extent that expert witnesses also assess justification memo quality in the same way, our results suggest that CPA firms should heed the manner with which justifications are written up, as well as their associated determinants. An important avenue for future research is to assess whether our findings related to how justification forms influence perceptions of memo quality apply to other groups such as expert witnesses and jurors.
APPENDIX  
Examples of Forms of Justification  
(Net Persuasive Evidence, Breadth of Issues, and Evidence Framing)

Panel A: Net Persuasive Evidence  
Preparers’ Conclusion: Require Additional Allowance for Doubtful Accounts  
Pro arguments:  
• Comparing the aging report of 1999 and 2000, 2000 sees an increase in the debts between 0–90 days.  

Con arguments:  
• Hired two experienced “credit agents” to monitor collectability.

Preparers’ Conclusion: Footnote Disclosure is Sufficient  
Pro arguments:  
• Client has implemented additional controls to deal with the change in credit policy.

Con arguments:  
• The credential of new debtors has decreased since they are smaller companies with lower credit ratings. Thus, risk of unrecoverability increases correspondingly.

Panel B: Breadth of Issues  
Internal control and related risk assessment  
• Company hired two experienced “credit agents” whose task is to monitor accounts receivable collections

Financial statement analysis:  
• Company’s aging accounts receivable for 91–120 days and greater than 120 days has decreased (2 percent).

Materiality  
• Sales made to new customers were not material to the total sales for the year.

GAAP  
• Based on the principle of prudence, suggest additional allowance.

Client-related factors:  
• Historically a good client for the last 5 years, hence, may have comfort with management representation.

Industry-related factors  
• Volatile nature of the industry.

Panel C: Framing Evidence to be Consistent with the Conclusion  
Preparers’ Conclusion: Require Additional Allowance for Doubtful Accounts  
• The company is very confident about its new strategy and feels that no additional accrual is necessary and merely a footnote is enough. However, based on the
principle of prudence and the volatile nature of the industry, suggest additional allowance.

**Preparers’ Conclusion: Footnote Disclosure is Sufficient**

- Although management adopted a more liberal credit policy by selling to smaller companies with lower credit ratings, management hired experienced credit agents to monitor accounts receivable collections, and, as such, no additional provision is required.

**REFERENCES**


