Does Auditing Deter Potentially Fraudulent Earnings Management?  
An Experimental Investigation

Qiu Chen*  
Doctoral Student

Khim Kelly#  
Assistant Professor

Steven E. Salterio* (corresponding author)  
Professor and PricewaterhouseCoopers/Tom O’Neal Faculty Fellow in Accounting  
Director, CA-Queen’s Centre for Governance  
ssalterio@business.queensu.ca  
(+1) 613-533-6926

*School of Business  
Queen’s University  
Kingston ON K7L 3N6

#School of Accounting and Finance  
University of Waterloo  
Waterloo ON N2L 3G1

(First draft January 18, 2009)  
March 13, 2011

Earlier versions of this paper were circulated under the titles “Do Audit Actions Consistent with Increased Auditor Scepticism Deter Potential Management Malfeasance?” and “Does Increased Auditor Scepticism Deter Management Malfeasance?” We appreciate the feedback of the Editor (David Cooper) and the reviewer for this Workshop as well as comments of colleagues at the International Symposium on Audit Research, CAAA Annual Meeting (especially the discussant Carolyn MacTavish), Auditing Section Mid-Year Meeting, Accounting, Behavior and Organizations Research Conference (especially two anonymous reviewers), attendees at the HEC Montréal, New England Behavioral Accounting Research Seminar, Mannheim University, University of Colorado at Boulder, University of North Texas, and York University accounting workshops, and those who have given us individual comments on earlier drafts of the working paper: especially Mohammad Abdolmohammadi, Jeff Cohen, Jonathan Farrar, Marlys Lipe, Alan Mayper, Naomi Soderstrom, Jay Thibodeau, Arnie Wright, and Mark Zimbelman.
Does Auditing Deter Potentially Fraudulent Earnings Management?
An Experimental Investigation

ABSTRACT

Audits are claimed to not only enhance the detection of fraud but also the deterrence of fraud. This study examines the ability of different audit procedures to deter potentially fraudulent earnings management and whether such different procedures influence managers’ perceptions about the ethicality of their intended earnings management. In an experiment with 171 senior corporate managers, we find that compared to the condition where the audit proceeds the same as last year, there is less intended earnings management when the nature of evidence collected changes towards increased probative value; and when the auditor conveys an attitude change via more critical inquiry combined with either an increase in the evidence extent (increased sample size) or a change in the nature of evidence. However, this reduction in intended earnings management is not found when the extent of evidence collected alone is increased. We also find, after controlling for the underlying ethical disposition of managers, that the different audit procedures affect managers’ perceptions of the ethicality of their intended actions. Interestingly, the conditions that engender greater intended earnings management also paradoxically increase managers’ sensitivity to the inappropriateness of their intended earnings management. Together these findings have implications for how different changes in audit testing may result in differential managerial responses about both the appropriateness of and the intention to commit fraud.

Keywords: potentially fraudulent earnings management; deterrence; effective audits; types of audit procedures; professional scepticism

Data Availability: Contact the third author.
Introduction

In recent years an increased emphasis has been put on the audit as not only a mechanism of fraud detection but also fraud deterrence, with the latter also known as fraud prevention (e.g., ISA 240; SAS 99; US Treasury Department 2008). As financial statement fraud often originates from a slippery slope of increasingly aggressive earnings management (Albrecht, Albrecht, Albrecht, & Zimbelman, 2009, p. 460-461), effective auditing should deter managers from starting on this slippery slope in the first place should in addition to its role in detecting overly aggressive earnings management. Auditors may not always detect fraud due to a variety of reasons including their inability to triangulate audit evidence (e.g., Harding & Wong, 2009; Trotman & Wright, 2009) and audit environmental factors (US Treasury Department 2008). Hence an understanding of how managers perceive different audit techniques and whether these techniques have different deterrence effects when managers are contemplating this slippery slope of aggressive earnings management is important.

The long-time claim of the financial audit as a fraud deterrence mechanism (AAA ASOBAC, 1973; Kranacher, 2006; Mautz & Sharaf, 1961; Panel on Audit Effectiveness (PAE) 2000; Wells, 2004) is more based on logical reasoning than on actual evidence (e.g., Schneider & Wilner, 1990; Uecker, Brief, & Kinney, 1981). The basic argument behind the deterrence claim is that managers report more honestly because they know their actions will be subject to audit scrutiny (Baiman, Evans, & Noel, 1987). If managers are averse to the discovery of their fraud, fraud deterrence should logically increase when the probability of fraud detection increases and people

---

1 Indeed, this assumption is a commonly accepted stylized fact employed to motivate many analytical models examining a variety of audit issues (e.g., Antle & Nalebuff, 1991; Fellingham & Newman, 1985).
perceive an increase in the probability of detection (Scheider, 2001). Our study employs the rigour of the experimental method to focus on the financial audit’s ability to deter fraud in a controlled setting where increased detection probability is minimized. If the audit deters potential fraud in this setting, the deterrence effect can only be stronger when coupled with increased probability of detection.

The relatively few research studies on the audit’s fraud deterrence effects have been limited by the choice of method, inconclusive results, and a focus on extreme experimental conditions (respectively, Zimbelman & Waller, 1999; Uecker et al., 1981; Schneider & Wilner, 1990). Zimbelman and Waller used an abstract experimental economics laboratory setting where undergraduate auditing student participants were assigned roles somewhat similar to “auditors” and “client managers”. In that experimental game, Zimbelman and Waller found weak evidence that “client managers” were less likely to report an asset at a misstated value when the “auditor” was less certain about the expected asset value that the “client manager” could report (see Bloomfield, 1999 for his discussion of the results in Zimbelman and Waller).

Uecker et al. (1981) found that exposing corporate managers, acting as CEOs, to a more “aggressive” auditor versus a less “aggressive” auditor, did not reduce managers’ potential to misstate receivables despite strong evidence that managers attended to the experimental manipulation of “aggressiveness”. Schneider and Wilner (1990) found, comparing extreme conditions of a “very competent and extremely aggressive” auditor versus no auditor at all, a decrease in the likelihood of financial reporting fraud only in the case that involved a highly material amount and a clear GAAP violation, but it had no effect whatsoever in the other two cases that were less extreme. Overall the evidence, while suggestive, provides little direct support for the deterrence effects of auditing and specific changes in audit procedures.

Our study examines the deterrence effects of specific and concrete changes in audit
procedures that could routinely occur in response to changes in audit risk: increasing the extent of the evidence collected and changing the nature of the evidence collected to have a higher probative value (ISA 240; ISA 330; SAS 99; SAS 110; Eilifsen & Messier, 2000; PCAOB 2008-006, A4-3).\(^2\) We also study whether a change in the auditor’s attitude towards the client managers, expressed via more critical inquiry, increases the hypothesized deterrence effects of a change in evidence extent or nature. Although auditing standards (ISA 330; ISA 500; SAS 110) regard inquiry as insufficient on its own to collect evidence of probative value, inquiry that conveys an auditor attitude change towards being more challenging of client management, in conjunction with other evidence changes, may result in added deterrence to client managers contemplating potentially fraudulent activities.

We conduct an experiment with a broad cross section of senior corporate managers who possess significant organizational experience such that undertaking the role of a middle to senior level corporate manager in the experiment would be natural. In the case, the participant is the manager of a division that is undergoing a less intensive audit than last year, who then learns from managers of other company divisions of their audit experiences this year. These experiences are manipulated as being the same as last year (SALY) audit, an increase in the extent of audit testing (i.e., an increase in sample size), a change in the nature of evidence collected (i.e., more external verification of management’s assertions), or a change in auditor’s attitude towards managers that is exhibited through more critical inquiry (i.e., more challenging auditor questioning) combined with each of the preceding three conditions.

We use two dependent variables in our analysis: a measure of intended earnings

---

\(^2\) We use the term “probative value” (i.e. having the quality or function of proving or demonstrating something (Davis & Follette 2002) and distinguish it from diagnosticity (i.e. diagnostic evidence supports one hypothesis more than another (e.g. Asare & Wright 1997)) as the client managers are not considering whether the new evidence as a result of these changed procedures will support one hypothesis versus another. Rather managers are assessing whether the change in audit procedures will result in proof of them committing fraudulent earnings management not just change the auditor’s assessed likelihood of the hypothesis of fraud/no fraud.
management that is potentially fraudulent (i.e., the amount of costs that the manager believes her peers would misallocate) and a measure of the manager’s perception of the ethicality of the intended earnings management (i.e., the difference between what the manager states she would misallocate and what she believes her peers would misallocate). We find that changing the nature of the audit evidence collected in the other divisions decreases managers’ intention to commit potentially fraudulent earnings management in their own division relative to a SALY audit and to an increase in the extent of evidence collected. Furthermore, when combined with an auditor attitude change that is conveyed by more critical questioning of managers, both changes in the extent of evidence collected as well as the nature decreases intended earnings management relative to a SALY audit. However, increasing the extent of the evidence collected alone does not reduce managers’ intention to commit potentially fraudulent earnings management in their own division vis-à-vis a SALY audit.

Furthermore, managers in the conditions where the audit was less effective in deterring the intended earnings management (i.e., SALY and extent increase alone) were also more sensitive to the (un)ethicality of their actions, even after controlling for their underlying ethical disposition. This increased sensitivity to the (un)ethicality of the intended earnings management is evidenced by a significantly larger gap between what managers state they would personally misallocate and what they expect their peers would misallocate in these conditions as compared to the smaller gap in the conditions where deterrence effect is stronger (i.e., nature alone, critical inquiry plus extent, critical inquiry plus nature). This suggests that situational factors (i.e., changes in audit actions) into their perceptions of ethicality even after controlling for their ethical disposition.

The remainder of this paper proceeds as follows. First, we develop our hypotheses based on organizational sense-making, interpersonal deception, and deterrence theories combined with our understanding of the institutional environment. We then outline our research design, our
experimental participants, the results of our experiment, and discuss the implications of these results for future research.

**Literature review and hypotheses development**

Our goal is to study the deterrence effect of auditing on managerial actions that are potentially fraudulent. Audit actions that increase deterrence generally coincide with actions that also increase the detection effectiveness of auditing. This is because the detection effectiveness of auditing may be one of the reasons why auditing deters fraud - people know that any perpetuated fraud has a higher chance of being discovered with auditing. Indeed, prior research that examines the deterrence effect of auditing often contrasts conditions with varying detection effectiveness, such as contrasting an “aggressive” auditor condition against a control condition of either no auditor at all (Schneider & Wilner, 1990) or a less “aggressive” auditor (Uecker et al., 1981). In Schneider and Wilner, the auditor versus no auditor setting clearly results in an increase in the likelihood of fraud detection, given even any auditor versus none at all is more likely to detect fraud. Even with this strong manipulation, Schneider and Wilner reported only a weak effect where an “aggressive” auditor compared to no auditor at all deterred potential fraud only in a setting that was characterized by an unambiguous GAAP violation involving a material amount. In the other two settings studied by Schneider and Wilner, no effect was found. Therefore, despite increasing the detection effectiveness of auditing, prior research provides very weak evidence of the deterrence effect of auditing.

We believe that examining whether auditing has deterrence effects when detection effectiveness is less likely to be an explanatory factor provides a stronger test of the audit’s deterrence effect. Hence, compared to prior studies, our study attempts to minimize the likelihood that actual detection effectiveness differs across the different auditing conditions that
we examine, so that we can better focus on the deterrence rather than detection abilities of different auditing conditions. Further, prior research does not examine specific auditor evidence collection activities or behaviours such that little insight can be given into what audit actions result in relatively more effective deterrence.

Uecker et al. (1981) featured relatively more and less aggressive auditors by manipulating two “aggressive” auditor actions (i.e., informing CEOs of their responsibility for the fairness of financial statements and requesting CEOs to include in their management representation their assessment of the reasonableness of an allowance for doubtful accounts). While participants strongly agreed with the characterization of these actions as being “more aggressive”, yet these “aggressive” auditor actions in combination had no deterrence effect. In Schneider and Wilner (1990), the “aggressive” auditor was only given a generic description (i.e., very competent, extremely aggressive, thorough, and strict). There was no information on what concrete changes in audit evidence gathering activities this aggressiveness entailed other than the vague notion that auditors were more “thorough in their work and strict about documentary evidence and substantiation of statements”.

In their experimental markets, Zimbelman and Waller (1999) did not control “auditor” reactions to “client manager” asset valuations but rather, by design, allowed auditors to choose the sample size they use to estimate the asset’s value, followed by another decision about whether they accept (or reject) the manager’s asset valuation. From their experiment, we learn that when auditors were not informed about the nature of the population from which the asset is randomly drawn and therefore were less certain about the asset’s expected value (as opposed to when the auditors knew the nature of the population), auditors increased their sample sizes and rejection rates of managers’ asset valuation, and managers are also simultaneously less likely to misstate the asset’s value. However, the experiment only allows us to conclude that the
manager’s reduced misstatement rate and the auditor’s increased sample sizes and larger rejection rates of manager’s proposed asset values co-occur but not whether they are causally related (see Bloomfield (1999) for further discussion of this point).³ Further, manager’s reduced misstatement rate in response to increased uncertainty faced by auditors is only weakly supported (Bloomfield, 1999). Thus, prior research gives us little guidance on what auditor actions are causally related to fraud deterrence.

Organizational sense-making, interpersonal deception theory and deterrence theory

We draw on three background literatures to understand better how the manager may react to changes in auditor actions and activities and be deterred from perpetrating fraud. First, we know from organizational behaviour research that managers attempt to make sense of changes that occur in their environment that might affect them, as a precursor to determining what their own behaviour will be (Weick, Sutcliffe, & Obstfeld, 2005). Hence, managers observing, or knowing through others of, changes in auditor behaviour will try to make sense of these changes prior to undertaking any actions in their own division, including potentially fraudulent earnings management. This sense making activity is consistent with experimental economics research in auditing that argues that auditors and managers behave strategically in anticipation of each others’ behaviour (e.g., Zimbelman & Waller, 1999).

Research on interpersonal deception indicates that all people (i.e., both those telling the truth and those attempting to deceive others) perceive increased suspicion of themselves by others when they learn of changes in the others’ described behaviour (Burgoon, Buller, Dillman, & Walther, 1995; Burgoon, Buller, Ebesu, White, & Rockwell, 1996). Therefore, changes in audit procedures toward those with a higher probative value should result in managers perceiving

---

³ Bloomfield (1999) in his discussion of this paper states “The complexity of the experimental setting hinders the interpretation of some of these results” (p. 164); further he highlights that it was particularly problematic in interpreting managers’ results (p. 163).
increased suspicion from the auditor. In addition, people attribute the others’ described behaviour change more to the others’ increased suspicion of them when they are lying or considering lying than when they are telling or planning to tell the truth (Buller, Strzyzewski, & Comstock, 1991; Burgoon et al., 1995). Hence, managers who are contemplating potentially fraudulent earnings management (i.e., a lie) are more likely to believe they are being suspected of something when they observe changes in audit procedures that may signal increased auditor suspicion.

Finally, deterrence theory (e.g., Becker, 1968) proposes three factors that affect people’s judgments about engaging in illegal or undesirable activities, namely certainty, severity, and swiftness of punishment. When people perceive an increase in the certainty of being caught in an illegal or socially undesirable act that results in severe and quick punishment, the costs of committing the act increases and hence the act’s expected utility decreases, thereby reducing the likelihood of people committing the act in the first place (Becker, 1968; Mendes & McDonald, 2001; Nagin & Pogarsky, 2001; Scheider, 2001). The deterrence effectiveness of punishment certainty is the one consistent empirical finding in the deterrence literature (Mendes & McDonald, 2001; Nagin & Pogarsky, 2001).4 Recent research also shows that it is the perceived rather than the objective probability of punishment that acts as a deterrent (Fischer, Wartick, & Mark, 1992; Lochner, 2007; Scheider, 2001). Furthermore, legal punishment may have less deterrence effect than people’s perceptions about whether a certain behaviour or action is wrong (i.e., their ethical disposition) (e.g., Foglia, 1997; Strelan & Boeckmann, 2006).

Changes in audit procedures

To be consistent with interpersonal deception theory, we need to determine what changes in audit procedures might suggest to a manager, contemplating a potentially fraudulent action, that

---

4 Research shows perceived punishment severity and swiftness have not been found to have a consistent significant effect on deterring criminal actions (Mendes & MacDonald, 2001; Nagin & Pogarsky, 2001).
the auditor’s suspicion has been aroused. The extant audit archival and experimental literature indicates that auditors can make a variety of changes to their audit program such as changing the timing, nature and/or extent of audit tests when they perceive changes in risk of material misstatement and their suspicion is raised (e.g., Bedard, Mock, & Wright, 1999).

Hence, we examine an increase in the extent of evidence sampled (i.e., expanding sample size for the same tests) and a change in the nature of evidence towards higher quality (i.e., from internal documents to external confirmations), that is, evidence of more probative value in assessing management’s assertions. Both these changes in audit procedures have been explicitly noted in recent audit standards as appropriate auditor responses to increases in risk of material misstatement (PCAOB 2008-006, A4-2-A4-6; ISA 240; ISA 330; SAS 99; SAS 110).

Hypotheses: Changing the extent or nature of audit evidence collected

Theoretically, as noted above, managers who are contemplating fraudulent actions (i.e., lying) would infer that the auditors are being more suspicious about the manager’s honesty when they perceive these changes in audit procedures, especially if the changes are towards procedures that increase the chance of fraud being detected. However, we focus on the manager’s perceptions as these changes occur in other company divisions while the changes in audit procedures in the manager’s own division hold constant, or likely reduce, the probability that fraud would be discovered in their division (a change towards substantive analytical procedures from tests of details). Hence, applying the above theoretical analyses we would predict:

H1 (SALY versus Extent). A focal manager made aware of an increase in the extent of audit evidence being collected in the other company divisions (Extent) is less likely to carry out potentially fraudulent earnings management in the manager’s own division compared to a

---

5 We do not wish to use the context of a new audit (i.e., versus no audit at all) since the audit is a fact of life for most large organizations. Further, it represents an extreme condition that is difficult to theoretically generalize from.

6 In our study we do not provide a rationale to our participants as to why the auditor has changed her procedures. We note that in many cases the auditor would not tell client management, especially those outside the C-suite, why the audit changes are being implemented. We examine two possible attributions that the managers might make for the changes, increased auditor suspicion about manager’s reporting honestly and changes to audit standards that require changes in audit procedures.

7 In our study, we ignore changes in timing because that would be difficult to meaningfully operationalize in an experimental context.
manager who knows that a “same as last year” audit is being conducted in the other company divisions (SALY).

**H2 (SALY versus Nature).** A focal manager made aware of a change in substantive nature of audit evidence collected towards more probative evidence in the other company divisions (Nature) is less likely to carry out potentially fraudulent earnings management in the manager’s own division compared to a manager who knows that a “same as last year” audit is being conducted in the other company divisions (SALY).

We next compare our “change in nature” and “increase in extent” manipulations to provide evidence on manager’s perceptions about the relative deterrence effects of different auditing changes. However, to form expectations about the relative deterrence effects of these two manipulations, further theorising is required as interpersonal deception theory in combination with perceived punishment certainty as developed to this point allows for no such distinction. Psychological research indicates that aversive events (e.g., seeing the auditors alter their behaviours in ways that could have negative consequences for the manager) that are more unpredictable are more stressful and create more sustained levels of anxiety (Baker & Stephenson, 2000; Grillon, Baas, Lissek, Smith, & Milstein, 2004). This added stress occurs because, unlike predictable aversive events, unpredictable aversive events do not provide people with periods of relative safety in which they are not under threat (Grillon et al., 2004) and do not allow people to prepare coping responses (Baker & Stephenson, 2000). Hence, it is possible that predictability of audit procedures provide another effect on perceived punishment certainty under deterrence theory.

In the context of our study, a manager contemplating fraud would experience higher levels of stress and anxiety with a more unpredictable auditor because unpredictability results in fewer “safety zones” where the manager can conduct fraud without being under the threat of unexpected discovery. A predictable auditor is psychologically less stressful for a manager considering fraud because managers, in their sense-making activities, may perceive that they could design their own fraudulent actions in such a way that already known audit procedures
would not detect the fraud (i.e., managers can prepare coping responses). Hence, if the extent of
the sample (i.e., sample size) is increased, but the same audit procedures and tests are employed,
managers are more able to anticipate where the auditors would look and hence ensure any fraud
the manager has committed would not be detected by those audit procedures and tests (Albrecht
et al., 2009; Ramos, 2003; Wells, 2004). As auditing standards note, “individuals within the
entity who are familiar with the audit procedures normally performed on engagements may be
more able to conceal fraudulent financial reporting” (e.g., ISA 240.69).

Therefore, auditing standards recommend that auditors introduce a certain degree of
unpredictability in their audit procedures when addressing fraud risks (e.g., ISA 240.66; SAS
99.50). Thus, if the auditor changes the nature of evidence collected, managers may perceive an
increase in the unpredictability of audit procedures (versus the increase in the extent of the same
audit evidence being collected with the same tests) that makes it more difficult to conceal a fraud,
thereby deterring the commitment of fraud in the first place (McKee, 2006). Again, we focus on
managerial perceptions as the changes of audit procedures that lead to more likely detection of
fraud, if committed, occur in other company divisions not our focal manager’s division where the
procedures change towards those less likely to detect fraud. Hence, we predict:

**H3 (Extent versus Nature).** A focal manager made aware of a change in the substantive nature
of audit evidence being collected towards more probative evidence in the other company
divisions (Nature) is less likely to carry out potentially fraudulent earnings management in the
manager’s own division compared to a manager made aware of an increase the extent of audit
evidence being collected (i.e., sample size increase only) in the other company divisions
(Extent).

_Hypotheses: Conveying a change of auditor attitude via more critical inquiry_

Research shows that changes in attitude generally affect a person’s behaviour or actions
(Eagly & Chaiken, 1993; see Petty, Wegner, & Fabrigar, 1997 for a review). One way the auditor
can visibly manifest to management a change in attitude is by making more pointed enquiries
and being more critical of management responses than in previous years (CICA, 2000; Liu, 2011).  

Following the research on interpersonal deception that focuses on how potential deceivers react to changes in environmental cues including changes in other’s attitudes (e.g., Burgoon et al., 1996; Burgoon et al., 1995), we expect that managers contemplating potentially fraudulent earnings management will be sensitive to auditors who become more critical in their questioning and challenge management assumptions and explanations more. Interpersonal deception research indicates that deceivers perceive more suspicion in the face of critical probing and therefore perceive a higher probability of lie detection (Buller et al., 1991; Burgoon et al., 1996). However, inquiry evidence alone may have limited probative value. After all, if people are considering committing a fraudulent action, it is quite likely they will be willing to lie about it in their responses to inquiries. Therefore, we advance no hypothesis about the effects of the change in auditor attitude exhibited through more critical questioning alone.

Combining critical inquiry and another change in audit procedure, however, could be mutually reinforcing in the sense that managers would perceive the change in attitude being reinforced by a change in actions that may actually uncover probative evidence of potential fraud as well. The combination of attitude and action changes compared to action change alone could signal to managers that the auditor is not merely going through the motions of changing audit procedures, but that the auditor would be reviewing the evidence obtained through the changes in audit procedures in more critical light. Further, the attitude change conveyed by more critical questioning combined with the audit action changes enables the manager to more readily explain the changes in the auditor’s activities in terms of the auditor having suspicions about the

---

8 In practice, auditors might exhibit attitude changes to clients in response to recent standards (e.g., SAS 109; ISA 315) that reinforce the requirement for the auditor to conduct an audit with professional scepticism, particularly when considering the risks of misstatement due to fraud. Both IAS Glossary of Terms and SAS 1 define professional scepticism as “an attitude that includes a questioning mind and a critical assessment of audit evidence”.

13
managers’ honesty. Hence, we posit that managers learning that auditors are more critical in their questioning in conjunction with a change in either extent or nature discussed previously will increase managers’ perceptions of auditor suspicion and the probability of detection more than the respective change in audit procedure alone. Therefore, our fourth hypothesis is as follows:

**H4a (Extent versus CI+Extent):** A focal manager made aware of an increase in the extent of audit evidence collected combined with more critical inquiry by the auditors in the other company divisions (CI+Extent) is less likely to carry out potentially fraudulent earnings management in the manager’s own division compared to a manager who knows of an increase in the extent of audit evidence collected alone in the other company divisions (Extent).

**H4b (Nature versus CI+Nature):** A focal manager made aware of a change in substantive nature of audit evidence collected towards more probative evidence combined with more critical inquiry by the auditors in the other company divisions (CI+Nature) is less likely to carry out potentially fraudulent earnings management in the manager’s own division compared to a manager who knows of a change in substantive nature of the audit evidence alone in the other company divisions (Nature).

Logically, if H1 (SALY versus Extent) or H2 (SALY versus Nature) is supported, then one would expect a replication of greater deterrence under CI+Extent than SALY and under CI+Nature than SALY.

**Research Question: Perception of ethicality of intended earnings management**

The perception of managers as to the ethicality of the intended earnings management is normally framed as a trait, that is, an internal person specific characteristic that influences an individual’s thoughts/behaviours consistently (e.g., Pervin, 1994, p. 108). The standard scales employed to measure an individual’s ethical orientation (e.g., the multi-dimensional ethics scale; see Cohen, Pant, & Sharp, 1996; Reidenbach & Robin, 1990) then are considered measures of the individual’s disposition along this ethical continuum. Indeed some research in other fields (e.g., drug testing as a mean of deterrence among athletes as in Strelan & Boeckmann, 2006) suggest that these ethical dispositions are the strongest factor in deterrence. However, situational constraints (i.e., auditor activity changes) may encourage or restrict the expression of traits in a given setting (Murphy, 1993; Trevino, 1986). For example, Murphy notes that traits such as
honesty may be more easily expressed in some situational contexts than in others. In the audit literature, the fraud triangle (SAS 99) also highlights the importance of both person specific characteristics and situational characteristics. Hence, in addition to examining how managers’ intentions to commit fraud differ with different auditor actions, we also consider whether these changes in audit procedures (i.e., changes in situational constraints) also differentially affect the managers’ perceptions of the ethicality of their response, over and above the effects of their ethical disposition.

We rely on research on social desirability for an appropriate approach to measure managers’ perceived ethicality of intended earnings management. Social desirability (often denoted as a bias) occurs when people under-report for themselves behaviours that they perceive as socially undesirable versus their report of those behaviours for their peers (Jo, 2000). People exhibit greater social desirability bias when they are responding to sensitive questions about behaviours they perceive to be more socially desirable/undesirable (Chung & Monroe, 2003; Fisher, 1993), and particularly so “among those who have something to hide” (Tourangeau & Yan, 2007, p. 863). Hence, a question about whether one intends to commit potentially fraudulent earnings management would prompt a more socially desirable response the more a person perceives the action as unethical, and the more the person intends to engage in that perceived unethical action and thus feel a stronger need to hide that intention. Therefore, we use the a measure of the social desirability in managers’ responses to measure their perceived ethicality of intended earnings management. We then expect that when auditing is less effective at discouraging people from engaging in potentially fraudulent activities (i.e., increased likelihood of earnings management) it also magnifies people’s need to distance themselves from the unethical action (e.g. Coughlan

---

9 “A question is sensitive when it asks for a socially undesirable answer, when it asks, in effect, that the respondent admit he or she has violated a social norm. Sensitivity in this sense is largely determined by the respondents' potential answers to the survey question; a question about voting is not sensitive for a respondent who voted” (Tourangeau & Yan, 2007, p. 860).
& Connolly, 2008). As a less effective deterrent creates a more “fraud-friendly” environment, people who plan to commit fraud in such an environment would feel a greater need to respond in a socially desirable manner to conceal their intentions to commit such an action.

Overall, this discussion leads to a similar set of predictions as developed in H1 to H4. However, the person-situation interactionist model of ethical decision making in organizations (Trevino, 1986) has found only modest support in the literature (Coughlan & Connelly, 2008; Paolillo & Vitell, 2002). Hence, while we have expectations based on our stronger theoretical development of H1 to H4 for this second construct, we are tentative enough about these expectations that we frame the issue in terms of the following research question:

**RQ:** Do the changes in situational constraints (i.e., changes in audit procedures) facing managers result in differences in managers’ perceptions of the ethicality of their intended earnings management consistent with the way they respond to the changes in audit procedures on their intention to commit earnings management as predicted in H1 to H4?

**Experimental design and method**

*Experimental design*

The experiment manipulates six between-subjects conditions (see Table 1 Panel A). The control condition describes this year’s audit testing in the other two divisions of a three-division company (i.e., not our focal manager’s division) as being the same as last year (SALY). The other five conditions describe the changes in audit procedures (vis-à-vis last year’s audit) occurring in the other two divisions (i.e., not our focal manager’s division) of a three-division company. The changes in the other two divisions are towards more probative value of evidence either by increasing the quantity of evidence collected (increased extent); changing the nature of the evidence collected (external confirmations versus reliance on internal documentation); or combinations of increased extent with more critical inquiry and change in nature with more critical inquiry. There is also a condition with a change in audit procedure for which there is no
hypothesis – change in auditor attitude alone via more critical questioning.

[Insert Table 1 Here]

In all conditions, the audit in the focal managers’ own division is less intensive than last year given the shift to substantive analytical procedures and less detailed substantive testing as part of the well known external auditor’s rotational audit approach. This is operationalized by a statement regarding the audit approach and by the manager receiving a memo from the Chief Financial Officer reminding them of the approach.

Participants

Participants were recruited from the MBA alumni who were at least five years post-graduation and EMBA alumni of a major Canadian business school that is ranked consistently high in the BusinessWeek rankings of MBA programs outside the USA.10 We enlisted prominent MBA faculty and administrators to send an invitation email directly to the alumni, with a follow-up email approximately one week later.11

Our study’s response rate, based on the 274 responses, is 12% of all working email addresses and 42% of click throughs (i.e., people who actually clicked through and went to the website). This response rate of 12% and 42% falls within the range of 7% to 44% rates documented in an analysis of e-mail and internet based surveys found in Schonlau, Fricker, and Elliott (2002).12 After excluding the 85 participants who failed our manipulation checks (described in the results section) and another 18 participants who omitted data critical to our analysis, our final sample

---

10 The five year post-MBA delay was used to ensure comparability with the more experienced EMBA participants. We did not keep track of MBA and EMBA participants separately given this control.
11 We followed carefully the literature on maximizing responses to email solicitations and web administered surveys (e.g., Manfreda & Vehovar, 2002; Schonlau, Fricker, & Elliott, 2002). Among the steps we took included having prominent faculty and administrators in the MBA and EMBA programs solicit participation after completing the case themselves. We also sent the emails at a time when the school’s alumni office was on a break from sending emails; hence the alumni are less likely to be suffering from over-contact by the school. We engaged an independent web design expert that reviewed our in-house IT department aided web design to ensure a user-friendly interface to encourage responses. Finally, we tested the case on recent MBA alumni (classes that were excluded from the study) to ensure that the web screens were user-friendly.
12 It is uncertain whether an invitation email has been read unless the participant clicks through to the website hosting the study (Sheehan, 2001). Hence, response rates are computed based on the percentage of emails sent out but not “bounced back” as well as the percentage of those who clicked through to the website.
size was 171 participants.\textsuperscript{13}

Table 2 presents demographic information about our final sample of 171 (there are no significant demographic or other differences with the sample of 248 which consists of participants with a complete data set but who failed one or more manipulation checks). Of particular note is the mean (13.9) and median (12.0) years of managerial experience.\textsuperscript{14} Randomization appears to have been successful as there were no significant demographic differences across experimental cells (all p’s $\geq 0.277$). However, we did find significant correlations between our primary dependent variables and two demographic variables (age and country); hence we include both of these variables as covariates in our analysis.\textsuperscript{15}

\textbf{Experimental procedures}

Participants clicked on the web-link provided in the invitation email that automatically took them to the website where our experimental instrument was located. Participants first read general information about the study, including all precautions being taken to ensure their anonymity, and clicked a consent button to participate in the study. Then they were randomly assigned by the survey software to one of the six conditions. The online software guided the participants through the managerial judgment case. Participants read the case and made their judgments (during which they could always access the case). Then participants answered post-experimental and demographic questions during which the case could not be consulted, as amongst the information elicited were recalls as well as interpretations of case facts.

\textsuperscript{13} As we report in our additional analysis section, results albeit weaker, replicate when all participants with complete data sets are included in our experimental analyses.

\textsuperscript{14} Note this is the lower bound of work experience as many of these managers have significant technical work experience prior to entering management ranks, a fact that this set of MBA programs sensitizes them to. Hence, they are unlikely to respond with anything other than actual “managerial experience” as asked for.

\textsuperscript{15} All other dimensions of participants’ demographic characteristic (e.g., gender, education) do not significantly correlate with the dependent variables. We also elicited answers to the question “how frequently have you experienced issues similar to the issue described in the case?” to determine the degree that participants had experience with our scenario. The mean response was at the mid-point of the scale anchored by “never” to “numerous occasions.” This variable is not significantly correlated with the dependent variables nor is it significantly different across experimental conditions.
Managerial judgment case

We adapted the case, with permission, from an earnings manipulation case developed by Cohen, Pant, and Sharp (2000) based on the cost tracking system of an actual company. The participant read about a project manager who worked in one of three divisions in a company, Pure Marine. The project manager was deciding how to account for $2 million of Research and Development (R&D) expenditure for internal reporting purposes. The $2 million ($2m hereafter) R&D expenditure was originally initiated and authorized for a nearly-completed Project K(3). Company policy required that all contract costs be assigned to the project for which they were initially incurred so as to reinforce the need for precise cost estimates at the time of the bid. However, the project manager was considering allocating the R&D expenditure to Project K(3) or to two other projects that would also likely benefit from the R&D expenditure but would not be completed until the next year. If the project manager allocated all of the $2m R&D expenditure to Project K(3), it would result in a loss of $1.2 million on the project and Pure Marine’s senior management viewed substantial losses on projects very unfavourably. Allocating at least $1.2m of R&D expenditure to the two other projects would avert reporting a loss on Project K(3) by moving the expenses to other projects that would be completed in the next year. Allocating at least $1.9m of R&D expenditure to the other projects would help the project manager to earn a bonus of up to $40,000 as Project K(3) would achieve its targeted profit of $700,000.

The case also links the cost allocation to external reporting and explicitly informs the participant that “As she sees it, the amounts in question are large enough such that it would mean that Pure Marine could actually meet or exceed analysts’ expectations for profitability this year.”

---

16 For external reporting purposes, R&D expenditures in a long-term contract can only be allocated to a contract if they are specifically reimbursable by the customer under both Canadian/International GAAP and US GAAP. To the extent that the R&D expenditures are not specifically reimbursable by the customer, allocating R&D expenditures to contracts (i.e., treating them as products costs that can be carried forward into future periods) for external reporting purposes would be a GAAP violation.
Further, “if the external auditors from Good and Better picks the project as one they investigate then it is likely that the allocation would be found. Given that the allocation would have potentially large implications for the financial statements it would likely be considered by Pure Marine management and the board of directors as a serious matter.” In the end, allocating any R&D expenditure to the two other projects violated company policy and could result in the manager fraudulently receiving an undeserved bonus, in addition to resulting in the company potentially issuing financial statements that have, at least, a qualitative materiality issue (e.g., meeting or beating analyst’s forecasts is one of the explicit examples given of qualitative materiality concerns in regulatory standards (i.e., SAB 99)).

The case takes place late in the fiscal year during an interim audit when substantive testing can be carried in the other divisions. This timing allows participants to make a potentially fraudulent accounting decision in the same year as the audit while knowing about the changes in audit procedures in the other divisions (or a SALLY audit in the control condition).

**Independent variables - changes in audit procedures**

As discussed previously, the changes in audit procedures employed in this case setting are based on our institutional understanding of how audit evidence can be changed in response to increased risks towards greater probative value and in addition, are consistent with approaches advocated by regulators (PCAOB 2008), existing professional standards (ISA 330; SAS 110), and practice as documented in audit textbooks (e.g., Knechel, Salterio, & Ballou, 2007, chap. 10). In addition, the reasonableness of these changes was discussed during the pre-test period separately with four audit partners who indicated agreement with our understanding. See Table 1 Panel A for all the conditions. The control condition was described to participants as follows:

---

17 Our managers accepted this allocation issue as being important for external reporting purposes. Their response to the question “If Sue allocated enough costs to the new projects such that K(3) meets its profit target, it will significantly increase the likelihood that Pure Marine will meet financial analysts’ expectations.” With -5 as “Strongly Disagree” and 5 as “Strongly Agree” resulted in 2.16 mean which is significantly greater (p<0.0001) then the scale midpoint of “0” and did not differ across conditions (p>0.50).
**Same as last year (SALY):** In the control SALY condition (Condition 1), participants are told that the audit in the other divisions is “going as smoothly as last year” with no changes in auditor behaviour compared to previous years (“auditing about the same number of projects”, “asking the same types of questions about the projects”, and “audit is going pretty much the same as it had in previous years”).

In the other conditions participants received information about the changes in auditor evidence collection activities described as follows:

**Increase in extent of evidence (Extent):** In the Extent condition (Condition 2), participants learned that the external audit team had “increased the number of projects audited by almost 50% over those audited in the previous years”, “the increase in number of projects audited placed more demands on support staff”, “the project managers are being asked more of the same types of questions as in the previous years and the volume of questions created more work for them”, and “the audit is more extensive than what it has been in previous years”. Hence, the nature of the audit evidence being examined in the other divisions is kept the same but the sample size is increased versus the prior year.

**Change in nature of evidence (Nature):** In the Nature condition (Condition 3), participants learned that the external audit team in the other two company divisions “had so completely changed their approach to selecting and auditing a project that Pure Marine staff could never anticipate what they would be asked for next”, “for the first time ever, the auditors are directly contacting external contractors and sub-contractors to verify work done on projects”, “the auditors are insisting that the contractors and sub-contractors provide written confirmation of work performed”, and “the audit is less predictable than what it has been in previous years”. Hence, the nature of the audit evidence collected in the other two divisions had shifted from evidence internal to the company to external confirmation as compared with prior years.
More critical inquiry (CI) combined with Extent and Nature: CI+SALY condition (Condition 4) is included as an additional condition in our research design as it provides a base for our manipulations of critical inquiry plus change in extent and critical inquiry plus change in nature. However, as stated in the theory section, we do not expect CI by itself to have a significant impact on intended earnings management, rather we include it for completeness of the design. To manipulate the attitude change of the auditors, managers were told that the external audit team “is being much more inquisitive than in prior years”, “the auditor’s questions seem more pointed”, “the auditors are not as accepting of the answers provided by the project managers as they had been in the previous years”, and the auditors were “being more critical in their questioning than usual”. This manipulation of critical inquiry followed guidance from the CICA Research Study on Audit Enquiry (CICA 2000) which is incorporated into the manuals and training programs of many audit firms.¹⁸

In the first combined experimental condition, CI+Extent (Condition 5), the auditor adopts more critical inquiry of management (as described in the CI+SALY condition) as well as increases the audit evidence sample size (as described in the Extent condition). In the second combined experimental condition, CI+Nature (Condition 6), the auditor is more critical with his/her inquiries of management (as described in the CI+SALY condition) and also changes the nature of the audit evidence collected (as described in the Nature condition). In all other respects the case description was the same as the SALY condition.

Ethics and this managerial judgement

Prior research employing the original version of this case showed that ethical dispositions of managers significantly affected the R&D cost allocation decisions made by managers (Cohen et al., 1996; Cohen, Pant, Sharp, & Holder-Webb, 2007). Hence, we also included the

¹⁸ One of the authors has carried out this training for two of these firms.
multi-dimensional ethics scale (MES) (Cohen et al., 1996; Reidenbach & Robin, 1990) in the post-experimental questionnaire so that individual ethical disposition can be controlled for (see Appendix for the MES used). In all conditions we elicited the MES measures for a manager allocating the entire $2m R&D expenditure to the two other projects (consistent with the approach of Cohen et al. 2007), providing a uniform anchor upon which to measure the ethical dispositions of the managers across control and treatment groups. Participants evaluated the action along a series of items that captured the moral equity dimension (e.g., just - unjust), the relativist dimension (e.g., traditionally unacceptable - traditionally acceptable), the utilitarian dimension (e.g., minimizes benefits while maximizes harm - maximizes benefits while minimizes harm), and the contractual dimension (e.g., does not violate an unwritten contract - violates an unwritten contract). We use these four multi-dimensional ethics sub-scales as control variables for individual differences in ethical judgment development in the main analysis.

**Dependent variables – allocations of $2m in Research & Development costs**

We elicit two allocation measures, the first is a measure of Self Allocation, that is, how much the participating manager would allocate (i.e., misallocate) to the two new projects. The second measure is Peer Allocation, which elicits how the participating manager believes his/her peers would make the same allocation. Table 1 Panel B reproduces both questions.

Studies show that for behaviours with moral overtones, people’s prediction of what others

---

19 Ethics researchers, including accounting ethics researchers, often advocate using the defining issues test (DIT) to control for differences in people’s ethical judgment development (e.g., Cohen et al., 2007). However, the on-line administration of the experiment combined with the length of time it takes to complete the DIT tasks made it impractical to administer it in this experiment. Our instrument already took on average 25 minutes to complete and most participants with incomplete data (nearly 10% of our potential participants) did not complete the sequence of items at the end of the instrument. Finally, Cohen et al. (2007) reported the MES replicated qualitatively their results obtained using the DIT for this particular case that we adapted.

20 We measured ethical disposition at the end rather than at the beginning of the case to avoid raising the ethical salience of matters in the case prior to measuring the dependent variables, although ideally we would have measured the ethical disposition first, as conceptually ethical dispositions are invariant person specific traits (Previn 1994). By anchoring the MES initial statement on the complete $2 million of R&D expenditures being allocated to the other projects; we endeavour to capture the ethical disposition of the managers by providing a common reference point. To the extent that participants incorporate situational constraints (i.e., audit procedures changes) when responding to the MES, we bias against finding any results for the dependent variable perceptions of ethicality.

21 We dropped the fifth mode of reasoning, the egoist dimension, because the two items measuring the egoist dimension did not clearly load on any factor in Cohen et al. (1996).
would do is more accurately correlated with their actual behaviour than people’s prediction of what they themselves would do (Epley & Dunning, 2000). The social desirability bias literature also indicates that indirect questions asking people to predict others’ behaviour suffer from less social desirability bias compared to direct questions asking people to predict their own behaviour (Fisher, 1993; Fisher & Tellis, 1998; Sherwood, 1981).

Therefore, we use the Peer Allocation measure as our dependent variable to proxy for managers’ intended earnings management. We control for Self Allocation in our analysis because prior research shows that responses for indirect questioning are affected by the responses for direct questioning (e.g., Jo, 2000, p. 140).

To measure the manager’s perception of the ethicality of their intended earnings management we employ the difference between the two measures (Peer Allocation less Self Allocation) as our second main dependent variable. Prior studies have measured social desirability bias as the difference between people’s report of what others would do and what they personally would do (Cohen et al., 2007; Chung & Monroe, 2003). People exhibit more social desirability bias when they perceive the behaviours in question as more socially undesirable (Chung & Monroe, 2003; Fisher, 1993; Tourangeau & Yan, 2007). Therefore, this measure allows us to report on whether managers incorporate situational factors (i.e. the changes in audit procedures) into their judgments of the ethicality of the intended earnings management while experimentally controlling for their underlying ethical disposition as measured by the MES.

---

22 Although our experiment had extensive upfront information about its anonymous nature and reminders of the same throughout the instrument, it does not completely mitigate the social desirability bias. As expected, a paired t-test indicated a significant difference between Self Allocation and Peer Allocation ($t=-6.55$, $p<0.001$), hence the need to carry out tests that reflected our a priori concerns for social desirability bias.

23 Cohen et al. (2007) used participants’ estimate of the likelihood that they personally would misallocate all $2m to the new projects as the dependent variable, and then employed as a social desirability bias control variable, the difference between participants’ estimates of the likelihood of them misallocating all $2m to the other projects and of their peers misallocating. Unlike Cohen et al. (2007), this measure of social desirability bias in our experiment (the difference in dollars between self and peer allocation) is significantly affected by our experimental conditions ($F=2.92$, $p=0.05$), and this is in fact what we examine in our research question. Hence, the use of this variable as a control variable would violate statistical assumptions and make the test results meaningless (Rosenbaum, 1984).
Pre-test of experimental instrument

In addition to the involvement of the four audit partners discussed above, the experimental instrument was pilot-tested with five graduating MBA students, 16 recent EMBA graduates from classes excluded from our email solicitation, and four faculty members who agreed to solicit participation from among the MBA and EMBA alumni. We used a talk aloud verbal protocol with the five graduating MBA students asking them to vocalize their thoughts as they read the experimental case to ensure that their understanding of the experimental instrument was congruent with what we intended. No material differences in understanding came to light in this process. We then had 16 recent EMBA graduates complete the on-line version of the instrument under the same conditions as experimental participants. The second pre-test lead us to revise our manipulations to make them more salient in the final instrument.

Results

Manipulation checks

We employed two manipulation checks to ensure participants understood their assigned condition with respect to nature and extent: (1) “Compared to prior years, you believe the external auditors have increased the number of projects audited in the other Groups” (i.e., manipulation check for an increase in extent of evidence), and (2) “Compared to prior years, you believe the external auditors changed the audit approach to require written confirmation from external contractors and sub-contractors” (i.e., manipulation check for a change in nature of evidence). Answers were elicited on a bipolar nine point scale (-4 “strong disagree” to 4 “strong agree” with 0 labelled as “uncertain”). Answering these manipulation check questions on the

---

24 Seven out of the 16 EMBA failed one or more of the manipulation checks. Due to a commitment about soliciting MBA participants electronically there was not the opportunity to carry out another pre-test of the revised instrument’s manipulations to determine if they were more salient.
correct side of the scale in the context of their assigned condition (i.e., “agree” in the conditions where extent/nature were manipulated and “disagree” in the conditions where extent/nature were not manipulated) was required for inclusion in the final sample. Out of our 274 responding managers, 31% (i.e., 85) failed either the extent or nature manipulation check.\footnote{Other than the SALY condition and the critical inquiry plus extent condition where manipulation check failure rate was near 20%, the other four conditions lost roughly an equal proportion of participants ($\chi^2=2.92$, $p>0.404$), near 37% on average. While one might expect that the SALY condition would have a lower manipulation check failure rate (given it is the simplest scenario), the reason for the lower manipulation check failure rate in the critical inquiry plus extent condition compared to the other conditions is not readily apparent.} We eliminated another 18 participants who lacked critical data we needed in order to carry out the statistical analysis, leaving a final sample of 171 participants who passed both manipulation checks and have a complete data set.

The manipulation checks for critical inquiry (“Compared to prior years, you believe the external auditors are not as accepting of the answers provided by project managers in the other Groups.” and “Compared to prior years, you believe the external auditors have become more critical in their questioning of project managers in the other Groups.”) indicated that all the five conditions (the four experimental and CI+SALY) viewed the auditors as being more critical with their questioning than the SALY condition (all $p$’s<0.001). This indicates that managers infer that any change in audit procedures is implicitly associated with more critical questioning, even when critical inquiry is not explicitly manipulated. We did not eliminate participants based on this manipulation check given that including participants who fail this manipulation check biases against finding results.\footnote{Practically speaking, it would result in eliminating so many participants in the non-critical inquiry conditions as to make statistical tests impossible to carry out.}

Participants also indicated in post experimental questions that they understood that allocating enough costs to the two new projects such that Project K(3) met its profit target would significantly increase the likelihood that the project manager would be eligible for a considerable
Further, participants generally believed that such an action is wrong as they believe that there would be negative consequences should it be discovered (financial consequences (mean=-1.80, median=-2), career consequences (mean=-2.33, median=-3), social standing within the company (mean=-1.78, median=-2), and social standing within the broader community (mean=-1.52, median=-1)). All four perceived consequences are significantly (p<0.001) negative (i.e., less than the scale’s “0” mid-point, on a scale of -4 for “very negative consequences” to 4 for “very positive consequences”).

Overall Tests of Hypotheses

Table 3 Panel A reports an omnibus tests of our hypotheses. The overall ANCOVA for intended earnings management (peer allocation) in Table 3 Panel A is significant (F=5.87, p<0.001), and the CONDITION variable is also significant (F=2.48 p<0.05). Table 4 Panel A indicates that the manager’s perception of ethicality of earnings management (Peer Allocation less Self Allocation) dependent variable is also significant (Condition, F=2.92, p<0.05) even after controlling for ethical disposition using the four components of the MES scale. As it is the predicted between-conditions differences that we are concerned about in our hypotheses, we follow Rosenthal, Rosnow, and Rubin (2000, p. 1-3) and carry out the planned tests.

[Insert Tables 3 and 4 and Figure 1 Here]

Individual Tests of Hypotheses and Research Question Results

Table 3 (Panels B and C) and Figure 1 present the descriptive statistics (i.e., least squares means adjusted for covariates and raw means) for each experimental condition as well as the specified comparisons based on our hypotheses. Testing H1 where focal managers learned of an increase in the extent of evidence (Extent) collected in the other divisions, we find no significant

---

27 In the final sample, over 96% percent of participants agreed with the statement “If Sue allocated enough costs to the new projects such that K(3) meets its profit target, it will significantly increase the likelihood that Sue will be eligible for a considerable bonus.” There were no differences between conditions (F = 0.97, p>0.435). Eliminating those participants who did not agree does not affect our analysis.

28 All p values reported are two-tailed unless otherwise specified.
reduction in cost allocation to other projects in the managers’ own division relative to the control condition where focal managers learned there was a SALY audit in the other divisions (p>0.50). Hence, H1 is not supported. Results from the research question parallel those found in H1, just as there were no significant differences in the intended amount of earnings management between SALY and Extent, managers’ perceived ethicality of the intended earnings management (Peer Allocation less Self Allocation) also did not differ between the two conditions (p > 0.50) as seen in Table 4 (Panels B and C).

Testing H2 where focal managers learned that the auditors changed the nature of the evidence (Nature) being collected (from mainly internal documents to external confirmations) in the other two divisions, we find a significantly lower cost allocation to the other projects in their own division than in the control condition (SALY audit). H2 is supported with a reduction in the Nature condition versus SALY of $400,909 in misallocation (one-tailed p<0.05, Table 3 Panels B and C). The research question results parallel the findings in H2. Perceived ethicality of the intended earnings management is lower in the Nature condition than in the SALY conditions (i.e., the gap between peer allocation and self allocation is reduced by $478,032, p<0.05, Table 4 Panels B and C). While managers intended to commit larger earnings management in SALY versus Nature, they also offered the more socially desirable response for themselves in the SALY condition. That is, they stated they personally would misallocate fewer dollars to the other projects whilst at the same time they stated their peers would commit a larger misstatement.

H3, which compares the Nature and Extent conditions, is also supported with a significant reduction of $489,538 in cost allocations to other projects (one-tailed p<0.005) for managers who learned of evidence nature changes in the other divisions versus managers who learned of extent changes in the same divisions. Again the results are consistent in the perceived ethicality measure as there is a significantly smaller gap between what managers said they would
misallocate as opposed to their peers in the Nature condition versus the Extent condition ($522,747, p<0.01, Table 4 Panels B and C). This latter finding again demonstrates that in the condition that exhibits more intended fraud (i.e., lower deterrence in Extent), managers claimed the ethical high road for themselves while stating their peers would commit a higher level of misstatement than in the condition with more effective deterrence (i.e., Nature).

H4a considers the effects of the auditor attitude change via more critical inquiry of management added to the Extent change compared to Extent alone. Supporting H4a, CI+Extent manipulation reduced the amount of cost misallocation versus Extent alone ($374,283, one-tailed p<0.05, Table 3 Panels B and C). Comparing the CI+Extent condition to the SALY condition, there is also a significant reduction in cost misallocation ($285,654, p<0.05, Table 3 Panels B and C) that is logically consistent with H1, albeit H1 was not supported. Thus, managers reacted to the mutually reinforcing changes exhibited by combining critical inquiry and increased sample size, but they did not react to expanding the sample size alone. Further, consistent with the H4a findings, there is also a significant reduction in perceived ethicality in CI+Extent versus Extent ($410,943, p<0.05, Table 4 Panels B and C), where again the condition that has less deterrence (i.e., Extent alone) also has a larger gap between what managers said their peers would do and what they would do.

H4b examines whether there is a similar effect when more critical inquiry is added to a change in the Nature of evidence collected. While CI+Nature is more effective than a SALY audit in reducing potentially fraudulent reporting ($291,261, one-tailed p<0.05, Table 3 Panels B and C) replicating H2, H4b is not supported. We found no incremental effect of an auditor’s attitude change via adding more critical inquiry to the Nature condition over Nature alone (two-tailed p > 0.50). Parallel results are reported for the manager’s perceived ethicality of the intended earnings management; no difference in perceived ethicality between CI+Nature and
Nature alone (two-tailed p > 0.50) whilst SALY has significantly higher perceived ethicality than CI+Nature ($416,496, p<0.05, Table 4 Panels B and C).

Finally, for completeness, Table 3 (Panels B and C) report the results found when critical questioning was added to the SALY condition (CI+SALY) versus the SALY alone, Extent alone and Nature alone conditions. In the three comparisons with the CI+SALY condition, while the effect was in the direction of reduced intended earnings management, it was not significant a finding consistent with our argument in the theory section. Similar results were obtained in the comparisons of perceived ethicality (see Table 4 Panels B and C).

Additional analyses

Including participants who failed the manipulation checks

We re-ran the analysis using the 248 managers (91% of total respondents) for whom we had a complete set of data (results not tabulated). Inclusion of all participants who failed the manipulation check biases against finding significant effects because at least some of the participants who failed the manipulation checks did not take the experiment seriously or did not understand their case. Even including all participants, all of our hypotheses test results replicated for the intended earnings management variable (peer allocation), albeit with slightly weaker levels of significance (all one tailed p’s < 0.065). The same results (with all two tailed p’s < 0.055) were also found for the perceived ethicality variable (Peer less Self Allocations) in our research question. This assures us that the reported results are not due to some systematic problem with the experimental instrument that resulted in manipulation check failures. The results are also stronger if we excluded the CI+SALY condition from all analyses for both the reduced (N=147) and full sample (N=208) set of managers.

Proportions instead of dollar amounts as dependent variable

While Uecker et al. (1981) focused their analyses on the amount of misreporting; they also
reported the proportion of people who misreported. We also examined the proportion of people who misreported using multivariate categorical ANOVA. 45 (26.3%) of participants reported that they personally would allocate costs to other projects versus 95 (55.6%) participants reported that their peers would do so. With the same control variables as in Table 3 Panel A, we found that using a dichotomous dependent variable (allocation to other projects versus no allocation) results in the same overall conclusion as using dollar amounts, i.e., a significant ($\chi^2(5)=11.23, p<0.05$) overall effect for Condition. Using categorical ANOVAs to test the individual hypotheses, all of our results for the amount of misreporting corresponded with our results for the proportion of people who misreported (results not tabulated).

We also analyze the proportion of participants who allocated at least enough costs (i.e., at least $1.2m) to the other projects to avert a loss on Project K(3), and the proportion of participants who allocated at least enough costs (i.e., at least $1.9m) to the other projects to earn a bonus. Again, we find that a smaller proportion of participants (26 or 15.2%) reported that they personally would allocate enough costs to avert a loss on Project K(3), than those (44 or 25.7%) who reported that their peers would do so. Similarly, only 17 (9.9%) participants indicated that they personally would allocate enough costs to earn a bonus, but 31 (18.1%) participants indicated that their peers would do so. While directionally replicating the previous results, the results using these two proportion dependent variables did not rise to conventional levels of significance.

*Support for interpersonal deception theory interpretation*

Based on interpersonal deception theory, we theorize that one of the key links in our hypotheses development is that managers would be sensitive to changes in audit procedures and associate those changes with increased auditor suspicion about the manager’s honesty. We asked all participants whether the auditor behaviours are due to increases in the auditor’s suspicion
about the company’s financial reporting honesty. Participants indicated their agreement on a 9 point scale with -4 being “strongly disagree” and 4 being “strongly agree”. Overall, managers attributed changes in audit procedures to the auditor’s suspicion of managerial honesty significantly more (all one-tailed p’s<0.05) in each of the four treatment conditions (Extent, Nature, CI+Extent, and CI+Nature) compared to the SALY condition. Further, as the magnitude of expenses the managers allocated to other projects in their own division increases (i.e., as the degree of potential fraud increases), managers’ attribution of auditor’s suspicion about managerial honesty also increases (t=2.48, one-tailed p<0.01).

We also gave managers a second reasonable attribution (per our pre-test elicitation of what could cause the auditor to change his or her actions), that is, the changes in audit procedures were due to changes in the regulatory environment. Providing this second attribution allows us to test if the suspicion attribution results reported above would be replicated in the same pattern with any reasonable attribution. We find that while the managers’ attribution of audit procedures to regulatory environment changes increases in each of the four treatment conditions compared to SALY (all one-tailed p’s<0.05), it is not related to the increases in cost misallocations in their own division (t=-0.10, p>0.90). Hence, these analyses provide support for our interpersonal deception theoretical analysis in developing our hypotheses.

Support for our deterrence theory interpretation

Based on deterrence theory, we posit that audit behaviour changes in the other divisions would increase perceived punishment certainty directly, or via a perceived increase in randomness of audit procedures, which then deters fraud. We asked managers six questions to assess their perceived punishment certainty: (1) how likely it is that the auditors will examine project files in the managers’ own division (scale of -4 (very unlikely) to 4 (very likely)); (2) how likely the misallocation would be discovered if the auditors examined the project file (same
scale); and (3) four questions on the consequences of the misallocation being discovered for the managers’ career, social standing within the company, social standing within the broader community, and personal finances (scales of -4 (very negative consequences) to 4 (very positive consequences)).

We find that managers perceived a greater likelihood of the auditors examining project files in their own division in each of the treatment conditions (with the exception of Nature) compared to the SALY condition, controlling for the same covariates as in Table 3 Panel A (all one-tail p’s <0.05, except one-tailed p < 0.105 for Nature versus SALY). This finding occurred even though the audit evidence collection changes were taking place in the other divisions, not the manager’s own division. In contrast, we expect and find no differences across conditions (F=0.54, p>0.50) when we asked managers “if the external auditors from Good and Better picks the project as one they investigate then it is likely that the allocation would be found.” Further, the responses were significantly greater than zero in each condition (all one-tailed p’s<0.001), indicating equally high perceived likelihood of discovery across all conditions. This question demonstrates that managers in all conditions equally understood that if a file were selected for audit in the manager’s own division then the misallocation to other projects would likely be discovered. Finally, employing the mean response to the four consequences of discovery questions, we expect and find no differences across conditions (F=0.97, p>0.44). The mean perceived consequences of discovery of the allocation to other projects were significantly negative (i.e., lead to punishment) in each condition (all one-tailed p’s<0.001). This result demonstrates that across all conditions managers understood the negative consequences of misallocation to other projects and perceived the negative consequences as being equally severe. This pattern of findings is consistent with punishment certainty theoretical analysis that motivated our predictions.
In our theoretical development for H3, we posited that relative predictability of the changes in auditor’s procedures would drive relative differences in deterrence between the Nature and Extent manipulations. Hence, we also asked managers to assess the perceived unpredictability of the auditor’s evidence gathering in each condition. We find that managers perceived greater unpredictability in the auditor’s actions in the Nature than Extent conditions (one-tailed p<0.001) and also in the CI+Nature than CI+Extent conditions (one-tailed p < 0.01). This result provides support for our theoretical analysis of auditor unpredictability employed to develop H3.

Discussion

The last two decades have seen a renaissance in the auditor’s focus on fraud detection and prevention given the importance that society attaches to the auditor’s role in reducing fraudulent financial reporting (e.g., the various expectations gap studies in many countries including the United States (AICPA, 1978; PAE, 2000; US Treasury Department, 2008), Canada (CICA, 1988), the United Kingdom (Auditing Research Foundation, 1989), and Australia (Australian Society of Certified Practising Accountants and The Institute of Chartered Accountants in Australia, 1994)). While professional standards have always enjoined the auditor to detect fraud that results in material misstatements (e.g., ISA 200; SAS 1), recent standards (e.g., ISA 315; SAS 109) further emphasize the need for the auditor to be particularly sensitive when considering the risks of material misstatement due to fraud. Specifically, SAS 99 and the newly revised ISA 240 (both fraud standards) are characterized by a dramatically increased emphasis on detection by adding new requirements for considering the ways companies can engage in fraud (Carpenter & Reimers, 2009) and adapting audit procedures accordingly when fraud risk is heightened.

29 Managers also perceive the auditor to be more unpredictable in each of the four treatment conditions (Extent, Nature, CI+Extent, and CI+Nature) compared to the SALY condition (all one-tailed p’s<0.001). However, adding more critical inquiry to either Extent or Nature does not significantly (p’s>0.30) increase perceived unpredictability (i.e., CI+Extent versus Extent and CI+Nature versus Nature).
(Carpenter, 2007; Marczewski & Akers, 2005). Despite these changes, regulators (e.g., PCAOB 2008-006, p. A4-2-A4-6) still express concerns that auditors are relying excessively on client explanations with insufficient scrutiny of the managers’ responses.

The extant audit archival and experimental audit literature indicates that auditors may respond to changes in audit risks (including fraud) in a variety of ways, ranging from retaining the audit plan from prior year without any adjustment to changing the nature and/or extent of audit tests (see Bedard et al., 1999 for a review of this literature). After making no adjustment to the audit plan, the most likely auditor response to heightened risk is an increase the extent of the evidence collected through increasing sample size (Bedard et al., 1999; Nelson, 2009). This is similarly shown in prior experimental studies where in response to increased risk, auditors’ change the extent of their audit tests by increasing the budgeted audit hours, but not the nature of their audit evidence (Glover, Prawitt, Schultz, & Zimbelman, 2003; Nelson, 2009; Zimbelman, 1997).

This portrayal of auditor response to increased risk in practice contrasted with how our managers react to these different types of auditor evidence changes is instructive. The most common auditor response to an increased risk (after making no changes) is for the auditor to increase the sample size for the same tests done last year. Yet our managers exhibited no increase in fraud deterrence from learning of an increased extent of audit testing alone in our experiment. While one could argue that this experimental finding is due to the changes in audit testing being described as carried out in the other divisions, this explanation does not account for the increase in deterrence when the auditor attitude change exhibited by more critical questioning is added to an increase in sample extent. However, research suggests that auditors are not likely to employ highly sceptical questioning towards their clients’ managers in the age of relationship management when auditors want to be “trusted business advisors” to their clients (McCracken,
Salterio, & Gibbins, 2008). This concern is reflected in ongoing regulatory statements about auditors not “sufficiently … challenge(ing) management’s forecasts, views, or representations” (PCAOB 2008-008, 20).

Our experimental results demonstrate that a change in the nature of audit evidence collected towards that of more probative value, even when it is not performed in the manager’s own division, leads to a substantial increase in deterrence. However, a change in the nature of the evidence collected is the least likely response of auditors to an increased risk of material misstatement found in practice or in the laboratory (Bedard et al., 1999; Nelson, 2009). Implicitly, it appears that auditors assume that the optimal audit procedures are being employed already, and that collecting and evaluating more of the same evidence is enough to reduce risk of material misstatement to a sufficiently low level (Fellingham, Newman, & Patterson, 1989; Fellingham & Newman, 1985; Mock & Wright, 1993). Nevertheless, a few studies have shown the auditor responds to risk increases with changes in the nature of the evidence collected but only in limited circumstances (Kizirian, Mayhew, & Sneathen, 2005; Mock & Turner, 2005; Shaub & Lawrence, 1996). Partially in response to this research along with their own regulatory inspection results, regulators and standard setters are requiring the inclusion of “elements of unpredictability in the selection of audit procedures to be performed” (e.g., PCAOB Auditing Standard 135c) so that the audit is not the same set of tests year in and year out. The very recent regulatory response is supported by our findings that change in nature of tests enhances the deterrence effects of the audit. However, our analysis of the attribution to professional standards changes indicates that there may be smaller incremental deterrence effects if auditor’s attribute these procedure changes to regulatory requirements rather than leaving the reason for changed procedures ambiguous. If relationship management is important to auditors as recent research indicates it is (e.g. McCracken et al 2009), it would seem likely that auditors would attribute these changes to the
changes in standards, potentially reducing the increases in deterrence that would occur if this attribution was not made.

**Conclusion**

The vast majority of effort in audit research has been focused on improving auditor detection capabilities with substantially less emphasis on the audit’s deterrence role (e.g., Hurtt, Eining & Plumlee, 2009; Glover et al., 2003; Knapp & Knapp, 2001; Mock & Turner, 2005; Phillips, 1999; Zimbelman, 1997). Nonetheless there is a strong claim made in the normative traditions of auditing about the audit’s deterrence and prevention role (e.g. AAA ASOBAC, 1973; Mautz & Sharaf, 1961) that continues to the present day (e.g., US Treasury Department 2008). Detection and deterrence are intimately interwoven in practice because an increase in the detection ability of the auditor, if it becomes widely known, should also lead to an increase in the deterrence ability of an audit. However, we know little about how audit changes affect managers’ underlying propensity to engage in potentially fraudulent action. Our study uses the advantage of the experimental method to study a scenario where increases in detection effectiveness of the audit action changes is minimal, allowing us to see how managers react to audit changes that have no direct effect on them, at least in the current period. Hence, relative to prior studies that examine deterrence where the detection effectiveness of the audit has also increased (Schneider & Wilner, 1990; Uecker et al., 1981; Zimbelman & Waller, 1999), our design reduces the possibility (albeit does not eliminate it) that any observed deterrence effect of audit action changes is attributable to an increase in the audit’s detection effectiveness. Overall, the goal of this research is to provide evidence on how different audit approaches deters potential fraud by minimizing the detection effects of audit approach changes.

We created a setting where a focal manager in one division learned that managers in the
same firm’s other divisions are facing either a same as last year audit or changes in auditor evidence collection behaviours versus the previous year. Across all conditions the managers are not and will not be subject to those changed procedures this year. Indeed, the focal manager’s division is actually receiving less audit attention, compared to the prior year, due to the well known external auditor’s rotational audit approach. Thus, given changes in audit procedures in this setting (i.e., the other divisions experience audit changes but not the division where the fraudulent action is being contemplated by the manager) indeed deter potentially fraudulent earnings management, it provides persuasive evidence that auditing has a deterrence effect on fraud merely by its effects on the perceptions of managers. In other words, the audit procedures (i.e., auditing) act as a deterrent even when the impact on detection effectiveness is limited.

Given concerns by regulators (US Treasury Department 2008) and standard setters (e.g. PCAOB 2007-001) that we do not know what audit actions are most likely to detect fraud in any given setting, an understanding of what procedures encourage managers not to commit fraud in the first place is valuable. Specifically, we find that managers who learn of a change in the nature of the audit evidence collected in the other divisions respond by reducing potentially fraudulent earnings management in their own division versus the same as last year condition and the increase in evidence sample extent only condition. Further, we find combining an auditor attitude change toward managers exhibited by more critical inquiry with either an increase in the extent of evidence collected or a change in the nature of evidence sought reduces intended earnings management compared to the same as last year audit condition. In contrast, an increase in the extent of audit evidence collected (or more critical inquiry alone) in the other divisions does not have a significant deterrence effect in the managers’ intention to commit potentially fraudulent earnings management relative to the same as last year audit condition.

We also find that the results for managers’ perceptions of ethicality of their intended
fraudulent behaviour mirror the results for their intended earnings management. Managers exhibit heightened perceptions of ethicality (even after controlling for underlying ethical dispositions) and thus perceive their intended earnings management to be more unethical in those conditions where they also intend to commit more earnings management. Paradoxically, when auditing is less effective at deterring earnings management, managers have a heightened sense of the inappropriateness of their intended earnings management and they respond in a more socially desirable manner reducing the amount of intended earnings management they will admit to but projecting a greater amount of such potentially fraudulent earnings management onto their peers.

Finally, our additional analysis shows that our theorizing based on interpersonal deception theory (i.e., the attribution of auditor behaviour changes to increased suspicion of managerial honesty) and deterrence theory (i.e., managers’ perceptions of punishment certainty and of the auditor’s unpredictability) are generally consistent with the results. Further, we find that it is the differences in perceived unpredictability of audit procedures that is associated with differences in fraud deterrence found between the Nature and Extent manipulations.

Our finding that changing the nature of evidence is a more effective deterrent than increasing the extent of evidence is particularly interesting in light of recent research on the effectiveness of triangulating audit evidence in detecting financial statement fraud. Bell, Peecher, and Solomon (2005) and Peecher, Schwartz, and Solomon (2007) have advocated that auditors triangulate audit evidence from both internal and external sources to identify inconsistencies that could improve the auditor’s ability to detect intentional misstatements. However, recent empirical evidence suggests that auditors are not very successful at triangulating evidence from external sources that contradicts with internal sources; even though the internal sources are more susceptible to manager’s intentional distortion (e.g., Harding & Wong, 2009; Trotman & Wright, 2009). Our research suggests that even if the detection effectiveness of auditors does not improve
with triangulation, the mere act of obtaining external sources of evidence may still have the benefit of deterring fraud if it increases perceived unpredictability in the auditor’s behaviour.

Of course, it is possible that the weaker deterrence found in the increase in extent condition vis-a-vis the change in nature condition may be attributed to the extent manipulation not being strong enough rather than inherent differences between the conditions. It is possible that increasing the sample size beyond that currently manipulated in our study (i.e., an increase in sample size of almost 50%) could have a stronger deterrence effect. However, our main objective is not to assess the relative deterrence merits of the changes in audit procedures as exactly operationalized in our study, rather we sought to use common auditor behaviour changes as advocated by professional standards and as found in practice, to test our theoretically derived hypotheses. Future research could examine whether there are inherent differences in changing the extent, nature, and timing of audit procedures that affect managerial perceptions and behaviours.

There are also other limitations to our research. First, we constructed a single period scenario so that we could observe directly the effects of changes in audit procedures on managers’ accounting behaviour in the current period. Ongoing deterrence by auditors in practice is likely a more complex phenomenon, but we believe we have captured essential elements for study. Second, we focus on senior managers in our research, but not a member of the C-suite (chief executive officer, chief financial officer, etc.) nor members of the accounting department. However, the Association of Certified Fraud Examiners reports that 30% of financial statement frauds reported are carried out by managers other than top executives and accountants (ACFE, 2008, p. 58). Further, it is widespread fraud deterrence, beyond the C-suite and the accounting department, which the auditing profession has frequently claimed as a positive effect of auditing (e.g., Kranacher, 2006; PAE, 2000; Wells, 2004) and that government and society expect (e.g.,
Lastly, it might be argued that the manipulation check failure rate is indicative of broader unspecified problems with our design. However, all of the hypotheses’ results are replicated when all participants are included in the analysis, a test that strongly biases against finding significant effects and hence suggests that the concern is not warranted.

There are at least three implications to be drawn from our research. First, we provide empirical evidence that demonstrates that changes in audit procedures have differential deterrence effect on managers even when increased detection effectiveness is unlikely to play a role. Hence, this means that manager’s perceptions of changes in audit evidence collection techniques can differentially deter potential fraud. When there is uncertainty about which audit evidence collection changes will result in improved fraud detection (e.g., US Treasury Department 2008) the results of this study can inform regulators, standard setters and auditors of what changes in auditor actions will deter fraud to a greater extent. Second, we show that various changes in audit procedures differentially affect managers’ perception of the ethicality of their intended earnings management even after controlling for their underlying ethical disposition. This finding reinforces our previous understanding that there is a person-situation ethical dynamic as presented in the fraud triangle (e.g., ISA 240, SAS 99, Wilks & Zimbelman, 2004). It suggests that situational factors can also include what the manager sees the auditor do, especially changes in evidence collection approaches. This is, we believe, a new finding to the fraud triangle literature (e.g., Wilks & Zimbelman, 2004). Third, we have provided a set of experimental results that suggest the deterrence effect becomes stronger with changes in audit procedures that increase manager’s perceptions of auditor unpredictability, an empirical finding that is consistent with what analytical models of auditing have concluded (e.g., Fellingham & Newman, 1985) but for which there existed little empirical evidence prior to our study.
References


American Institute of Certified Public Accountants (AICPA). *Statements on Auditing Standards (SAS) No. 1*, “Due professional care in the performance of work”.


American Institute of Certified Public Accountants (AICPA). *Statements on Auditing Standards (SAS) No. 109*, “Understanding the entity and its environment and assessing the risks of material misstatement”.

American Institute of Certified Public Accountants (AICPA). *Statements on Auditing Standards (SAS) No. 110*, “Performing audit procedures in response to assessed risks and evaluating the audit evidence obtained”.


Bell, T. B., Peecher, M. E., & Solomon, I. (2005). The strategic-systems approach to auditing. IN T. Bell, & I. Solomon (Eds.), *Cases in strategic systems auditing: KPMG and University of Illinois at Urbana-Champaign Business Measurement Case Development and Research Program* (pp. 1-34). KPMG LLP.


International Auditing and Assurance Standards Board (IAASB). *International standards of auditing (ISA) No. 200*. “Objective and general principles governing an audit of financial statements”.

International Auditing and Assurance Standards Board (IAASB). *International standards of auditing (ISA) No. 240*. “The auditor’s responsibility to consider fraud in an audit of financial statements”.

International Auditing and Assurance Standards Board (IAASB). *International standards of auditing (ISA) No. 315*. “Understanding the entity and its environment and assessing the risks of material misstatement”.

International Auditing and Assurance Standards Board (IAASB). *International standards of auditing (ISA) No. 320*. “Materiality in the identification and evaluation of misstatements”.


Rosenbaum, P. R. (1984). The consequences of adjustment for a concomitant variable that has been affected by the treatment. *Journal of Royal Statistical Society Series A (General), 147* (5), 656-666.


Notes:

**LS Mean Peer Allocation:** The least squares adjusted mean response to the question “If members of your peer group (e.g. fellow alumni) were making the allocation decision, how much of the $2 million Technical Improvement R&D expenditure do you think they would likely charge to the new projects?” adjusted for the four multi-dimensional ethics sub-scales, Age, Country, and Self Allocation to Other Projects (per Table 3 Panel A) and tabulated in Table 3 Panel B.

**Raw Mean Peer Allocation:** The unadjusted mean responses to the above question on Peer Allocation.

**SALY:** Same as last year audit in the other divisions

**Extent:** Increase in the extent of evidence (i.e., increase in sample size) in the other divisions

**Nature:** Change in the nature of evidence (i.e., from internal documents to external confirmations) in the other divisions

**CI+Extent, CI+Nature:** Auditor attitude change as exhibited by more critical inquiry (CI) combined with Extent or Nature conditions, respectively.
TABLE 1
Experimental Design

Panel A: Between-Subjects Independent Variable Manipulations

<table>
<thead>
<tr>
<th>Nature of Management Inquiry</th>
<th>Same as last year</th>
<th>Increase in audit evidence sample size</th>
<th>Change in nature of audit evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same as last year</td>
<td>Condition 1</td>
<td>Condition 2</td>
<td>Condition 3</td>
</tr>
<tr>
<td>SALK</td>
<td>Extent</td>
<td>Nature</td>
<td></td>
</tr>
<tr>
<td>More critical inquiry</td>
<td>Condition 4</td>
<td>Condition 5</td>
<td>Condition 6</td>
</tr>
<tr>
<td>CI + SALK</td>
<td>CI + Extent</td>
<td>CI + Nature</td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Dependent Variable Questions a

1. Self Allocation

1. As Sue Davies, how much of the $2 million Technical Improvement R&D expenditure would you likely charge to the new projects? (the unit "m" represents "millions")

<table>
<thead>
<tr>
<th>$0</th>
<th>$0.25m</th>
<th>$0.5m</th>
<th>$0.75m</th>
<th>$1.0m</th>
<th>$1.25m</th>
<th>$1.5m</th>
<th>$1.75m</th>
<th>$2.0m</th>
</tr>
</thead>
</table>

Dollar allocation: $__________

2. Peer Allocation

2. If members of your peer group (e.g. fellow alumni) were making the allocation decision, how much of the $2 million Technical Improvement R&D expenditure do you think they would likely charge to the new projects? (the unit "m" represents "millions")

<table>
<thead>
<tr>
<th>$0</th>
<th>$0.25m</th>
<th>$0.5m</th>
<th>$0.75m</th>
<th>$1.0m</th>
<th>$1.25m</th>
<th>$1.5m</th>
<th>$1.75m</th>
<th>$2.0m</th>
</tr>
</thead>
</table>

Dollar allocation: $__________

Note:
a We use the Peer Allocation as our dependent variable to proxy for intended earnings management (see Table 3), and the difference between Peer Allocation and Self Allocation as our dependent variable to proxy for social desirability bias or perceived ethicality of intended earnings management (see Table 4)
### TABLE 2
Manager Participant Profile (N = 171<sup>a</sup>)

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
</tr>
<tr>
<td>26–40</td>
<td>37.4%</td>
</tr>
<tr>
<td>41–50</td>
<td>43.9%</td>
</tr>
<tr>
<td>51–60</td>
<td>17.5%</td>
</tr>
<tr>
<td>61–70</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>over 70</td>
<td>&lt;1%</td>
</tr>
<tr>
<td><strong>Years of managerial experience</strong></td>
<td></td>
</tr>
<tr>
<td>Mean = 13.89 years</td>
<td></td>
</tr>
<tr>
<td>Std. dev. = 7.83 years</td>
<td></td>
</tr>
<tr>
<td>Upper Quartile = 20 years</td>
<td></td>
</tr>
<tr>
<td>Median = 12 years</td>
<td></td>
</tr>
<tr>
<td>Lower Quartile = 8 years</td>
<td></td>
</tr>
<tr>
<td><strong>Gender (% female)</strong></td>
<td>14.7%</td>
</tr>
<tr>
<td><strong>Canada as country of birth</strong></td>
<td>75.4%</td>
</tr>
<tr>
<td><strong>Education level</strong></td>
<td>100% MBA&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Educational major</strong></td>
<td></td>
</tr>
<tr>
<td>Accounting</td>
<td>4.76%</td>
</tr>
<tr>
<td>Other business</td>
<td>54.76%</td>
</tr>
<tr>
<td>Non business</td>
<td>40.48%</td>
</tr>
<tr>
<td><strong>Accounting Designation</strong></td>
<td>14.3%</td>
</tr>
<tr>
<td><strong>Lawyer or attorney</strong></td>
<td>2.4%</td>
</tr>
<tr>
<td><strong>Employment profile</strong></td>
<td></td>
</tr>
<tr>
<td>Public company</td>
<td>42.4%</td>
</tr>
<tr>
<td>Large private company</td>
<td>9.1%</td>
</tr>
<tr>
<td>Medium or small private company</td>
<td>27.3%</td>
</tr>
<tr>
<td>Government or not-for-profit organization</td>
<td>17.0%</td>
</tr>
<tr>
<td>Other</td>
<td>4.2%</td>
</tr>
</tbody>
</table>

**Notes:**
<sup>a</sup>Excludes those who failed the manipulation checks, however, there is no significant difference between the profiles of those participants who passed the manipulation checks and those who did not.

<sup>b</sup>Low female recruitment in technology-oriented full-time MBA program and executive MBA programs.

<sup>c</sup>Consistent with our drawing from an MBA e-mail list.
### TABLE 3

**Intended Earnings Management:**
**Amount Peers Would Allocate To New Projects (Peer Allocation)**

**Panel A: Analysis of Covariance (ANCOVA) of Intended Earnings Management (Peer Allocation)**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares (^a)</th>
<th>Mean Square (^a)</th>
<th>F Value</th>
<th>2-tailed p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition (^b)</td>
<td>5</td>
<td>5.01</td>
<td>1.00</td>
<td>2.48</td>
<td>0.034</td>
</tr>
<tr>
<td>Moral equity dimension (^c)</td>
<td>1</td>
<td>0.64</td>
<td>0.64</td>
<td>1.60</td>
<td>0.208</td>
</tr>
<tr>
<td>Relativist dimension (^c)</td>
<td>1</td>
<td>0.05</td>
<td>0.05</td>
<td>0.13</td>
<td>0.719</td>
</tr>
<tr>
<td>Utilitarian dimension (^c)</td>
<td>1</td>
<td>0.05</td>
<td>0.05</td>
<td>0.12</td>
<td>0.728</td>
</tr>
<tr>
<td>Contractual dimension (^c)</td>
<td>1</td>
<td>1.35</td>
<td>1.35</td>
<td>3.34</td>
<td>0.070</td>
</tr>
<tr>
<td>Age (^d)</td>
<td>4</td>
<td>2.93</td>
<td>0.73</td>
<td>1.82</td>
<td>0.128</td>
</tr>
<tr>
<td>Country (^d)</td>
<td>1</td>
<td>1.49</td>
<td>1.49</td>
<td>3.69</td>
<td>0.057</td>
</tr>
<tr>
<td>Self allocation – other projects (^e)</td>
<td>1</td>
<td>20.40</td>
<td>20.40</td>
<td>50.56</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Error</td>
<td>155</td>
<td>62.55</td>
<td>0.40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
\(^a\) All sum of squares and mean squares are in \(\text{‘000,000,000,000}\).
\(^b\) Condition is the experimental treatment as described in Table 1 Panel A.
\(^c\) There are four multi-dimensional ethics sub-scales that measure the ethical judgment development of participants; see questions in the Appendix.
\(^d\) Age and Country are as measured and reported in Table 2.
\(^e\) Self allocation: amount participant reported they personally would misallocate to other projects.

**Panel B: Least Squares Means and Raw Means (Standard Deviations) of Intended Earnings Management (Peer Allocation)**

<table>
<thead>
<tr>
<th>Condition (^a)</th>
<th>N</th>
<th>LS Mean (^b)</th>
<th>Mean (Std. Dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same as last year (SALY)</td>
<td>38</td>
<td>$1,064,769</td>
<td>$710,526 (778,786)</td>
</tr>
<tr>
<td>Extent</td>
<td>32</td>
<td>$1,153,398</td>
<td>$926,250 (773,512)</td>
</tr>
<tr>
<td>Nature</td>
<td>24</td>
<td>$663,860</td>
<td>$568,750 (784,609)</td>
</tr>
<tr>
<td>Critical Inquiry (CI) + Extent</td>
<td>29</td>
<td>$779,115</td>
<td>$583,448 (762,151)</td>
</tr>
<tr>
<td>CI + Nature</td>
<td>24</td>
<td>$773,508</td>
<td>$665,000 (724,065)</td>
</tr>
<tr>
<td>CI + SALY</td>
<td>24</td>
<td>$862,655</td>
<td>$646,250 (723,991)</td>
</tr>
</tbody>
</table>

Note:
\(^a\) All conditions are defined in Table 1 Panel A.
\(^b\) LS Mean: Least squares mean adjusted for other variables in ANCOVA (four multi-dimensional ethics sub-scales, Age, Country, and Self allocation).
Panel C: P-values Associated with Pair-wise Comparisons of Least Squares Means of Intended Earnings Management (Peer Allocation)\(^a\)

<table>
<thead>
<tr>
<th>Condition</th>
<th>SALY</th>
<th>Extent</th>
<th>Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent</td>
<td></td>
<td>H1</td>
<td>H2</td>
</tr>
<tr>
<td>Nature</td>
<td>0.011</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>CI + Extent</td>
<td>0.039(^b)</td>
<td>0.012</td>
<td>0.523 (2-tailed)</td>
</tr>
<tr>
<td>H1 replicated</td>
<td></td>
<td>H4a</td>
<td>H4b</td>
</tr>
<tr>
<td>CI + Nature</td>
<td>0.045(^b)</td>
<td>0.032 (2-tailed)</td>
<td>0.562 (2-tailed)</td>
</tr>
<tr>
<td>CI + SALY</td>
<td>0.25 (2-tailed)</td>
<td>0.10 (2-tailed)</td>
<td>0.29 (2-tailed)</td>
</tr>
</tbody>
</table>

Note:
\(^a\) All p-values for hypotheses are one-tailed, unless otherwise stated, in keeping with the directional nature of hypotheses. Two-tailed p-values are reported for H1 and H4b because the direction of the effect is opposite of that predicted in the hypotheses. P-values for non-hypothesized comparisons are two-tailed.
\(^b\) One tailed tests as they are tests replicating the directional predictions of H1 and H2.
### TABLE 4

**Perception of Ethicality of Intended Earnings Management**  
*(Peer Allocation less Self Allocation)*

**Panel A: Analysis of Covariance (ANCOVA) of Perception of Ethicality (Peer Allocation Less Self Allocation)\(^a\)**

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares(^b)</th>
<th>Mean Square(^b)</th>
<th>F Value</th>
<th>2-tailed p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>5</td>
<td>6.94</td>
<td>1.39</td>
<td>2.92</td>
<td>0.015</td>
</tr>
<tr>
<td>Moral equity dimension</td>
<td>1</td>
<td>1.46</td>
<td>1.46</td>
<td>3.08</td>
<td>0.081</td>
</tr>
<tr>
<td>Relativist dimension</td>
<td>1</td>
<td>0.10</td>
<td>0.10</td>
<td>0.21</td>
<td>0.647</td>
</tr>
<tr>
<td>Utilitarian dimension</td>
<td>1</td>
<td>0.002</td>
<td>0.002</td>
<td>0.00</td>
<td>0.955</td>
</tr>
<tr>
<td>Contractual dimension</td>
<td>1</td>
<td>1.07</td>
<td>1.07</td>
<td>2.26</td>
<td>0.135</td>
</tr>
<tr>
<td>Age(^d)</td>
<td>4</td>
<td>1.76</td>
<td>0.44</td>
<td>0.93</td>
<td>0.450</td>
</tr>
<tr>
<td>Country</td>
<td>1</td>
<td>1.48</td>
<td>1.48</td>
<td>3.12</td>
<td>0.079</td>
</tr>
<tr>
<td>Error</td>
<td>156</td>
<td>74.07</td>
<td>0.47</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
\(^a\) The dependent variable is the difference between Peer Allocation and Self Allocation. All other variables are defined in Table 3.
\(^b\) All sums of squares and mean squares are in ‘000,000,000,000.


<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>LS Mean(^b)</th>
<th>Mean (Std. Dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same as last year (SALY)</td>
<td>38</td>
<td>$596,910</td>
<td>$500,000 (660,160)</td>
</tr>
<tr>
<td>Extent</td>
<td>32</td>
<td>$641,625</td>
<td>$637,188 (841,413)</td>
</tr>
<tr>
<td>Nature</td>
<td>24</td>
<td>$118,878</td>
<td>$104,167 (729,490)</td>
</tr>
<tr>
<td>Critical Inquiry (CI) + Extent</td>
<td>29</td>
<td>$230,682</td>
<td>$181,034 (490,520)</td>
</tr>
<tr>
<td>CI + Nature</td>
<td>24</td>
<td>$180,414</td>
<td>$175,417 (556,409)</td>
</tr>
<tr>
<td>CI + SALY</td>
<td>24</td>
<td>$373,604</td>
<td>$406,667 (825,478)</td>
</tr>
</tbody>
</table>

Notes:
\(^a\) All conditions as defined in Table 1 Panel A.
\(^b\) LS Mean: Least squares mean adjusted for other variables in ANCOVA (four multi-dimensional ethics sub-scales, Age, and Country).
Panel C: P-values Associated with Pair-wise Comparisons of LS Means of Perceived Ethicality (Peer Allocation Less Self Allocation)\(^a\)

<table>
<thead>
<tr>
<th>Condition</th>
<th>SALY</th>
<th>Extent</th>
<th>Nature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent</td>
<td>0.795</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nature</td>
<td>0.012</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>CI + Extent</td>
<td>0.037</td>
<td>0.023</td>
<td>0.568</td>
</tr>
<tr>
<td>CI + Nature</td>
<td>0.024</td>
<td>0.016</td>
<td>0.764</td>
</tr>
</tbody>
</table>

Note: \(^a\) We employ the directional predictions from H1 to H4 in the text regarding research question results. However, as this is a research question, all reported p-values are two-tailed.
APPENDIX – Ethics Questionnaire
(Explanations in italics were not in the participant version and the items were randomly ordered)

Ignoring your own answer to the case, suppose that Sue Davies charged the entire $2 million Technical Improvement R&D expenditure to the new projects. Please indicate your views about this action on the following scales by choosing a value on the scale closest to the term that best describes your reaction to the statement at the top.

Charging the $2 million to the new projects is:

**Moral equity dimension**

<table>
<thead>
<tr>
<th>Moral equity dimension</th>
<th>Unjust</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Just</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morally right</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not morally right</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acceptable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unacceptable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Relativist dimension**

<table>
<thead>
<tr>
<th>Relativist dimension</th>
<th>Culturally unacceptable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culturally acceptable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traditionally acceptable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Utilitarian dimension**

<table>
<thead>
<tr>
<th>Utilitarian dimension</th>
<th>Produces the least utility</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Produces the greatest utility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimizes benefits while maximizes harm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Contractual dimension**

<table>
<thead>
<tr>
<th>Contractual dimension</th>
<th>Violates an unwritten contract</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does not violate an unwritten contract</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Violates an unspoken promise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The probability that I would undertake the same action is:

<table>
<thead>
<tr>
<th>The probability that I would undertake the same action</th>
<th>High</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The probability that members of my peer group (e.g. fellow alumni) would undertake the same action is:

<table>
<thead>
<tr>
<th>The probability that members of my peer group would undertake the same action</th>
<th>High</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ethical (recoded such that 8 is ethical and 1 is unethical)-Used in analysis as control variable

The action described above is:

<table>
<thead>
<tr>
<th>The action described above is:</th>
<th>Unethical</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>