Students as Surrogates for Managers: Evidence from a Replicated Experiment

Kim Trottier

Beedie School of Business

Simon Fraser University

Burnaby, BC, Canada

V5A 1S6

kim_trottier@sfu.ca

and

Irene M. Gordon[a]

Beedie School of Business

Simon Fraser University

Burnaby, BC, Canada

V5A 1S6

Irene_Gordon@sfu.ca

[a] Corresponding author email: Kim_Trottier@sfu.ca

We are grateful to the Beedie School of Business students who participated in this research and to Stephen Spector for his cooperation in contacting his students. We also thank Karel Hrazdil and Craig Emby for comments on earlier drafts. Irene Gordon acknowledges the support of the CGA-BC Association and Kim Trottier acknowledges research funding supplied by the Beedie School of Business and Simon Fraser University's President's Research Grant.

Students as Surrogates for Managers: Evidence from a Replicated Experiment

ABSTRACT

Using students as surrogates for managers in experiments is commonplace, yet this practice is not always valid. To explore when the use of student samples is appropriate, we replicate an experiment previously conducted employing a sample of senior managers involved in financial reporting. The result is that although student and manager responses are significantly different from a statistical perspective, both samples lead to the same conclusion for this experiment. The findings suggest that having some disassociation between students and the target population they are meant to represent does not necessarily make them inappropriate surrogates. To examine when inferences are best supported, we explore the comparability for student sub-groups and managers.

Keywords: Students; surrogates; experiments; ethics; knowledge/expertise.

Students as Surrogates for Managers: Evidence from a Replicated Experiment

Researchers conducting experiments often use students as "convenience samples" under the assumption that responses obtained are representative of the researcher's target population. The appropriateness of generalizing student results has been "formally recognized, empirically examined, and heatedly debated in a variety of disciplines for more than five decades" (Peterson, 2001, p. 450). Fuchs and Sarstedt (2010) point out that the tacit acceptance of student samples is incongruent with the deficiency in evidence that supports its use. Given the continued propensity to use students as surrogates to derive inferences regarding managers, it is important to further explore the validity of this practice. Our paper addresses the question: "When can students validly be used as surrogates for managers and why are they valid surrogates in some conditions but not in others?"

Three important issues are raised by the literature examining whether students are appropriate surrogates in business experiments. First, when replicating an experiment with students, the data often show that students respond differently from managers to some but not all of the questions in the experiments, a finding that is difficult to interpret (e.g., Alpert, 1967; Abdel-Khalik, 1974; Ashton, 1974; Ashton & Kramer, 1980; Bean & D'Aquila, 2003). Second, the p-values can be inconsistent across test statistics (Abdel-Khalik, 1974; Ashton & Kramer, 1980), and are susceptible to a myriad of methodological problems (Lindsay, 1995). Third, when interpreting research the tendency is to focus on whether results are statistically significant (Lindsay, 1995). However, establishing whether observed differences alter the *inference* and *conclusions* of the experiment, rather than focusing on the statistical properties of the responses, seems a more logical way of evaluating surrogacy, since the use of students is only problematic if it leads to incorrect conclusions. This is the approach used in our analysis.

Regardless of our results, we realize the reality is that we are not likely to ever reach the sweeping conclusion that students are *always* or *never* suitable surrogates for managers. Therefore research needs to eschew such generalizations, and move towards an understanding of specific student characteristics that improve the correspondence between student and manager response.

Our paper addresses these issues by analyzing a financial reporting experiment conducted on a sample of senior managers involved in their firms' accounting decisions (Trottier, 2013),[1] and subsequently replicated with a sample of undergraduate business students recruited from a fourth year Accounting Theory course. The experiment presented a setting where a fictitious manager must choose whether to record an impairment loss on an asset. In the scenario there was some uncertainty about the true current value of the asset as well as a small probability the value would recover in the next period, which motivated the manager to omit the loss from the financial statements even though the accounting rule required its recognition. The experiment employed a 2x2 factorial design to determine if the decision to record the impairment was influenced by accounting rules and/or by the existence of a bonus plan.

We examine the responses in three ways. First, we compare responses across the student and manager samples to explore the appropriateness of surrogacy from the perspective that is consistent with prior research. Second, we explore whether inferences are the same, even in the presence of statistically significant differences in responses. Following suggestions of previous research, we complete the analysis by identifying subgroups of students that are more appropriate surrogates, as well as student characteristics that may confound results.

Our paper contributes to the literature by placing emphasis on the inferences that may be derived from student surrogate samples rather than whether responses are similar from a statistical perspective. Additionally, we inform the literature by indicating when the composition

of student samples may cause researchers to be more cautious in interpreting their results. Finally, finding similar inferences from a replicated experiment strengthens the results from the manager paper by showing these are repeatable even under differing conditions (i.e., change in sample) (Lindsay & Ehrenberg 1993).

The remainder of this paper is presented in four sections. The next section outlines the prior literature. The following section develops our hypotheses about whether student and manager responses should differ. The subsequent section explains our research methods, and is followed by the presentation and discussion of our results. This is followed by additional analysis on how certain subsets of students compare to managers. We conclude our paper and discuss limitations of our study in the final section.

2. Prior literature

The use of students as surrogates for managers has been a controversial issue for decades (Dickhaut, Livingstone & Watson, 1972; Ashton & Kramer, 1980; Greenberg, 1987; Peterson, 2001). In a review of the highly ranked literature in marketing and business management, Fuchs and Sarstedt (2010) find that two-thirds of experiments published between 2005 and 2007 use students in spite of the mixed evidence on the validity of this method. The practice of using students in business research experiments is likely to continue due to costs, accessibility, students' willingness to participate and potential overuse of professionals (*e.g.*, Sears, 1986; Shuptrine, 1975; Simon, 1979; Enis, Cox, & Stafford, 1972; Libby, Bloomfield & Nelson, 2002). The justification that business students are future managers (Remus, 1986) provides another reason to use them.

Much of the literature on students as surrogates explores either attitude (opinions) or cognitive (decision making) differences between the two groups (e.g., Osgood, Suci, & Tannenbaum, 1957; Copeland, Francia, & Strawser, 1973; Berkowitz & Donnerstein, 1982;

Hughes & Gibson, 1991; Houghton & Hronsky, 1993). Peterson (2001, p. 450) suggests that rigorous analysis is needed to avoid positions "based on conjecture or anecdotal evidence." Therefore our literature review is focused on papers that make use of statistical analysis to assess the similarity of student and manager responses in experiments.

Four empirical studies using structured experiments explicitly examine student attitudes/perceptions compared to other groups. Examining the generation of arguments to support a subordinate's firing, Alpert (1967) finds students to be poor substitutes for managers due to differences in experience. Copeland, Francia, & Strawser (1973) compare accounting student and accountant attitudes toward financial reporting practices, finding students to be poor proxies. Ennis, Cox & Stafford (1972) explore race/country of origin bias while Shuptrine (1975) probes whether female education graduate students respond similarly to other female consumers. These last two studies provide inconclusive results, suggesting life experience plays an important role as to when students are better surrogates. In examining whether students respond to advertisements in the same way as other consumers, James & Sonner (2001) find students are good proxies when the two groups' ages are similar.

The evidence is also mixed in experiments simultaneously conducted with students and managers to gain insight into whether they respond similarly in decision-making settings. Abdel-khalik (1974) examines 40 lending cases using MBA students compared to managers and finds students are inadequate surrogates for managers, while Ashton and Kramer (1980) use 32 internal control judgment cases and find upper division auditing students to be adequate surrogates for auditors. These two studies explore multiple differences and similarities at a very granular level hindering their ability to reach an overall conclusion.

In a structured production scheduling exercise, Remus (1986) finds students to be good proxies when their educational background is similar to managers as do Elliott, Hodge, Kennedy

and Pronk (2007) for complex integrative investment decisions. Other papers using structured approaches, such as Hughes and Gibson (1991) with a training experiment, Fehr and List (2004) with a trust experiment and Chang and Ho (2004) with a project continuation experiment, also find students unsuitable substitutes for experienced managers.

The accounting and auditing domain yields inconclusive evidence as to the suitability of student or inexperienced accounting surrogates. For example, Houghton and Hronsky (1993, pp. 142-3) use an experiment that examines the structure (meanings/definitions) and placement (quantification) of accounting concepts. They find students are adequate surrogates for practitioners in the task of recalling meanings or definitions but not for the "more precise task" of quantifying those meanings. Using financial reporting cases with ethical considerations, Bean and D'Aquila (2003) find accounting students unsuitable surrogates while Liyanarachchi and Milne (2005) (environmental liabilities disclosure experiment) and Mortensen, Fisher, and Wines (2012) (a classification judgment experiment) conclude accounting students are good surrogates. Auditing researchers have examined whether experienced and inexperienced auditors respond differently to situations (Weber, 1980; Libby & Frederick, 1990; Frederick, 1991, Nelson, 1993; Tubbs 1992; Hoffman, Joe, & Moser 2003). In structured experiments experienced auditors are found to show better understanding, plan better and make better judgments compared to inexperienced students.

Table 1 provides an overview of studies that specifically assess whether students make good surrogates. Our table indicates that depending on context, the evidence is mixed.

[Insert Table 1 here]

This research primarily rests on students' age, education or work experience. Given the overall inconclusive results, there is a need for more research on the use of students as

surrogates. We look to research studies beyond those exploring students as surrogates to provide additional support for studying specific student characteristics.

Grades have been used in previous education related studies (e.g., Globiel & Phillips, 2001; Kohli, Peng, & Mittal, 2011). Findings from such studies indicate students with higher grades are better able to demonstrate and use their knowledge, especially in more challenging settings. Vera-Munoz, Kinney, & Bonner (2001) examine the effect of knowledge (and types of knowledge) on managers' and accountants' ability to perform tasks, finding knowledge affects outcomes. Based on these two sets of findings, we examine grades as a means of capturing student knowledge.

With respect to culture, evidence indicates that cultural differences affect personal traits. In particular, culture has been found to influence people's level of uncertainty avoidance and assertiveness (König, Steinmetz, Frese, Rauch, & Wang, 2007) and the importance of personal achievement vs. communal harmony (Kitayama, Park, Sevincer, Karasawa, & Uskul, 2009). Cultural differences are also found in ethical views (Berger & Malinowski, 2004; Gift, Gift & Zheng, 2012), response to ethical dilemmas (Bernardi, Witek, & Melton 2009), willingness to report peers' questionable acts (Brunton & Eweje, 2010), decision making (McCabe, Dukerich & Dutton, 1993), and work-life value systems (Zhang, Straub & Kusyk, 2007). Given that our study examines a financial reporting issue related to ethics, we examine how students' responses are influenced by their cultural backgrounds.

Finally, gender differences have been studied. Halpren (2012) indicates that men and women overall show no differences in intelligence, but differences do exist on tests of reasoning/spatial ability (edge goes to men) and verbal ability (edge goes to women). Other studies (e.g., Byrnes, Miller, & Schafer, 1999; Olsen & Cox, 2001) find that women tend to be more risk averse than their male counterparts. The ethics-based literature has found gender

differences with females being more moral (Bernardi & Arnold 1997), less tolerant of unethical behaviour (Lopez, Echner, & Olson-Buchanan, 2005), inclined to be ethical (Albaum & Peterson, 2006), more concerned with ethical responsibilities (Atakan, Burnaz, & Topcu, 2008) and more ethically aware when making ethical judgments (Eweje & Brunton, 2010). This literature supports exploration of whether gender affects how managers and students respond to our case.

3. Study background, framework and hypothesis development

A survey was created to investigate truthful reporting as a result of changes in accounting standards. The project was motivated by the adoption of IAS No. 36, which permits reversals of impairment losses on long-lived assets, and was intended to explore whether this standard would improve reporting not only when an asset recovers its value, but also through more truthful reporting when the asset is initially impaired. Survey details are provided in the appendix.

The survey began with a scenario involving a fictitious manager (Fred) who identifies an existing but somewhat subjective asset impairment. Participants were asked to assess the likelihood that Fred would record the impairment, where each participant received one of four scenarios: Fred either had or did not have a bonus plan; and the impairment was either permanent or reversible. Respondents were asked whether they believed Fred would record the impairment, rather than whether *they personally* would record it. This format is used to reduce the social desirability bias (Fisher, 1993; Chung & Monroe 2003) – a tendency for respondents to provide answers they believe would be viewed favourably – but creates other issues examined in Section 5.3.

Our analysis is based on the student and manager responses to the survey which was administered to both groups within a short time frame. Managers were identified through the

Global Insight database as high-level individuals involved in their firms' financial reporting, and students were recruited after having completed a fourth year accounting theory course.

The responses received from the manager sample have been analyzed and published (Trottier, 2013), warranting a summary of those findings and their implications. The author finds that both the reversibility and bonus plan have an effect on the perceived likelihood the impairment will be recorded. Having a bonus plan leads to a statistically significant decrease in the likelihood it will be reported, but allowing the impairment to be reversed if the value of the asset recovers results in an increase in the likelihood of truthful reporting to a level that eliminates reporting differences between bonus motivated and non-bonus motivated managers. The managers' responses are consistent with agency theory. In an agency theory context, managers are assumed to take the course of action that increases their expected utility (Jensen & Meckling, 1976; Watts & Zimmerman, 1986). Compensation plans, including bonuses, are meant to align managers' interests with those of shareholders and other stakeholders (Smith & Watts, 1982; Lambert & Larcker, 1987; Brozovsky & Sopariwala, 1995; Duru, Mansi, & Reeb, 2005), but they may also induce a manager to adopt accounting treatments that increase his/her bonus (e.g., Healy, 1985; Holthausen, Larcker, & Sloan, 1995).

While we assume managers were able to draw on their work experiences (and possibly education) to identify when the case subject is likely to choose an accounting treatment resulting in a higher bonus, the students in this study have less direct experience with the effects of compensation plans. However, all students in the study have had courses that include exposure to agency theory, bonus plans and discussion of how these interact and influence managers' accounting choices. Therefore we expect students will recognize that the manager's decision to record the asset impairment will affect the manager's utility. A priori, we have no reason to expect students to respond differently from the managers surveyed in Trottier (2013).

The existing literature also predicts no difference in responses, since the survey involves decision-making (as in Remus, 1986), and the knowledge of the students is similar to managers' (as in Mortensen et al., 2012). This leads to our main hypothesis (stated in alternative form):

HYPOTHESIS 1: Students will assess the same likelihood of impairment as managers across all treatment conditions.

Section 5.2 explores other dimensions that could inform which subset of students make decisions most similar to managers, namely work experience, culture, grades, and gender.

4. Method

Manager and student participants were asked to evaluate whether Fred would record an asset impairment he has identified. There is uncertainty about the degree of impairment, but his subjective assessment indicates that the current value is less than the recorded amount, resulting in a large overstatement of book value. Accounting rules in the case specify that impairments should be recognized if the book value is overstated, which would decrease net income. However there is a small chance the asset will recover its value in the next few years. The participant's estimate of the likelihood Fred would record the impairment is analyzed with two way between-subjects factors: bonus plan (with bonus, without bonus) and type of impairment (reversible, permanent). The four versions of the instrument were randomly assigned to participants in each of the manager and student samples. After reading their scenario, participants were asked to evaluate the likelihood that Fred would record the impairment, on a scale of one (very unlikely) to seven (very likely), and then answer a series of demographic questions that provided a basis for exploring attributes associated with their responses.[2]

4.1 Participants

As noted in Trottier (2013) the manager sample was generated by sending an email invitation to managers identified in a database obtained from Global Insight, a service that provides contact

information on Canadian managers. A total of 118 Canadian managers completed the experiment, for a response rate of 7.8 percent, which approaches other Internet surveys response rates for managers (e.g., Graham, Harvey, & Rajgopal, 2005). The primary titles/positions of this 118-manager sample are CEO (21 percent), General Manager (25 percent) or Owner (53 percent) with a majority of managers coming from the manufacturing (43 percent) and wholesale (48 percent) industries. The preamble to the survey stated the case would be based on financial statements, which we assume deterred participants with no accounting knowledge. The majority of manager respondents reported having responsibility for financial statement preparation (82 percent) or estimates used in those statements (a further 8 percent).

The student sample was obtained by sending an email invitation to business undergraduate students who had completed the fourth year Accounting Theory course at a Canadian university (Simon Fraser University) in the previous three years. A total of 649 students were invited to participate. A set of 103 students completed the instrument for a 16% response rate. The majority of these respondents had between one and three school terms left at the time they took the survey, with approximately 15% having recently completed university. The relatively low number of graduates is reasonable given that the course they were recruited from can be taken earlier than in fourth year, and many students work part-time or participate in the co-operative education program, effectively extending the duration of their studies.

5. Results

5.1 Differences between students and managers

Descriptive statistics on participants are reported in Table 2. Based on the 649 students invited to participate, Panel B shows there is no significant difference between the percentage of day students (16.63%) and evening students (12.82%) who completed the instrument. However as

reported in Panel C, a significantly larger proportion of "A" students completed the experiment (27.75%) than did "B" (13.74%) and "C" students (6.94%).[3]

[Insert Table 2 here]

Table 3 provides summary measures of both student and manager responses regarding the main question in the experiment. Participants were asked to provide a score, on a scale of one (very unlikely) to seven (very likely), reflecting their perceived likelihood the scenario's manager would record a loss impairment. Panel A of Table 3 shows that managers evaluated a higher likelihood that the loss would be recorded (with a mean score of 4.85) than the students (with a mean score of 4.37), with a difference that is significant at the 10% level (p-value = .07), the conventional threshold for a nondirectional hypothesis. This initial result does not support Hypothesis 1, which presumes no difference between the two samples. Panel B reports the mean likelihood score by treatment condition. Results show that scenarios with reversible impairments (with and without bonus) generated a higher likelihood assessment that impairments would be recorded for both the student and manager samples. In both samples, the lowest likelihood score was obtained in the scenario where Fred had a bonus plan *and* the impairment was permanent.

[Insert Table 3 here]

A test of difference in means by treatment combination (Panel B) shows the "permanent impairment with bonus" category was evaluated significantly lower by students than managers (p=0.08), and "permanent impairment with no bonus" category was different at a level that is almost significant (p=0.12). It appears the students and managers diverged in their perception of how the permanent impairment accounting rules would affect Fred's choice.

The differences by treatment are explored more precisely in Table 4. Here we estimate the differential and interactive effects with an ANCOVA model for consistency with Trottier (2013) and to allow for interpretation of the coefficients.[4] The model is:

$$Likelihood = \alpha + \beta_1 Reversible + \beta_2 NoBonus + \beta_3 Interaction + \delta_1 SDum + \delta_2 SDum*Reversible + \delta_3 SDum*NoBonus + \delta_4 SDum*Interaction + \varepsilon \quad [1]$$

where *Likelihood* is the participant's assessment (from one to seven) of the likelihood the impairment would be recorded, *Reversible* is a variable equal to one if the subject's scenario treated impairments as reversible, and *NoBonus* is a variable equal to one if the scenario did not include a bonus plan. The *Interaction* variable is equal to one if the scenario had both conditions: reversible impairment with no bonus plan. The model is then augmented with variables that capture differences in student responses. The *SDum* variable is an indicator variable equal to 1 for student participants. This indicator variable is multiplied by the three main variables (*Reversible, NoBonus,* and *Interaction*) to construct a set of variables designed to capture differences in student responses to the treatment conditions (*SDum*Reversible, SDum*NoBonus,* and *SDum*Interaction*). The result from estimating Equation 1 with the combined manager and student samples is reported in Panel A of Table 4. The statistically significant parameter estimates are interpreted as follows.

The coefficients suggest truthful reporting is more likely if impairments are reversible (β_1 = 1.58) or Fred is not incentivized by a bonus plan (β_2 = 1.14). Removing both deterrents to truthful reporting increases the likelihood the asset impairment will be recorded, but the combined effect is moderated by the interaction term (β_3 = -1.16). The significant coefficient on *SDum* (δ_1= -1.29) tells us students rated the likelihood of truthful reporting lower than managers overall, but this discrepancy between the two groups was less pronounced in scenarios with reversible impairments (δ_2 = 0.95). Hence the data do not support Hypothesis 1, but are instead consistent with the branch of existing research that concludes students provide systematically different responses from managers in experimental settings.

[Insert Table 4 here]

If the aim of this paper was to determine whether students' and managers' responses were different at a statistically significant level, the above test would be sufficient; we would conclude that students are *not* appropriate surrogates for managers. Ultimately, however, we think the question of surrogacy lies in whether it leads to the same inference rather than similar responses. We establish this by estimating the ANCOVA model separately for the student and manager sample, and comparing the inference. After eliminating the student interaction terms, the model is reduced to:

$$Likelihood = \alpha + \beta_1 Reversible + \beta_2 NoBonus + \beta_3 Interaction + \varepsilon \quad [2]$$

Panel B of Table 4 tabulates the results from an ANCOVA estimate of Equation 2 for students and managers separately. The first column reports the coefficients obtained with the manager sample, which are consistent with the original study on the likelihood of impairments (Trottier, 2013). The coefficients from estimating the model with the student sample are tabulated in the second set of columns in Panel B. While there are differences in magnitudes between the regression coefficients obtained with the student sample and the manager sample, especially regarding the *Reversible* indicator variable (with a value of 2.05 for managers and 1.76 for students), the *inference is identical*. Both samples yield results that indicate the fictitious manager in the scenario is more likely to record a loss if it is reversible or if he has no bonus plan, with an interaction term that attenuates the combined effect.

Consequently, although there are significant differences between student and manager responses (causing us to reject Hypothesis 1), the inference and conclusions from the experiment are the same. Although students were overall more skeptical, both sets of participants believe the bonus plan and irreversibility reduce the likelihood the impairment will be recorded. Hence using students for this particular accounting experiment would have been appropriate.

5.2 Additional analysis

A second objective of the present research is to explore whether the external validity of our experiment is improved by utilizing certain subsets of students. External validity is key to assessing whether results are generalizeable. The existing literature suggests certain characteristics of the experiments and specific student traits could aid in improving external validity (e.g., Ashton & Kramer, 1980; Houghton & Kronsky, 1993; Sears, 1986). In order to advance improvements in the use of convenience samples, this paper includes an assessment of the similarity of student and manager responses for several characteristics.

The instrument's demographic section asked students to provide information we use to examine if some sub-groups of students responded more similarly to managers. The demographic section of the instrument included questions on the students' country of birth (and country where they lived their first 15 years), gender, age, and work experience so we could explore whether these dimensions lead to systematic differences in responses. The grade and course section (i.e., day/evening) information was available from the student database where the sample was drawn, and allows us to further examine the notion of experimental realism.

For this analysis, we apply a simple test of difference in means, to avoid potential multicollinearity and loss of power from estimating a regression model that includes all the demographic variables. Results are corroborated with a more detailed test (not tabulated) that examines differences between student and manager responses within treatment conditions. Specifically, for each student observation we calculate the absolute difference between the student's likelihood score and the mean (also median) manager score that was obtained in the student's treatment condition, and explore whether student characteristics are associated with smaller or larger absolute differences. Results from our detailed test are consistent with the test of difference in means tabulated in Table 5.[5]

[Insert Table 5 here]

5.2.1 Work experience

Druckman and Kam (2011) propose the validity of an experiment improves with "mundane realism", that is the extent to which the scenario is likely to occur in the normal course of the subjects' lives. Students with work experience should experience a higher level of "mundane realism" when participating in the experiment, therefore their responses should be similar to the managers. Given the findings of previous studies (e.g., Hoffman et al. 2003), we expect that students with work experience are more likely to respond similarly to managers due to mundane realism.

To test this, we divide the students into three categories based on work experience. Panel A of Table 5 indicates that students with fewer than five years of work experience provided responses that were significantly lower than managers, but students with five years or more of work experience had responses similar to managers. These findings lead to the conclusion that students with work experience are better surrogates for business managers in experiments.

The analysis in Panel B of Table 5 corroborates our result on work experience. We test whether responses from students taking evening classes were more similar to managers' responses compared to students in day classes. This provides a more objective measure than our groupings in Panel A, but relies on the assumption that working students are more likely to take evening classes. The results tabulated in Panel B of Table 5 are consistent with the work experience findings - evening student responses more closely match managers than do day student responses.

5.2.2 Knowledge

We use additional personal attributes to tease out which students represent better proxies for managers (Bonner, 2008) from a knowledge/intelligence perspective, captured through education

(proxied by student's course grade). The related literature (e.g., Globiel & Phillips, 2001; Kohli et al. 2011 and e.g., Vera-Munoz et al. 2001) leads us to expect that students who possess more knowledge will respond more similarly to managers than students who have less knowledge.

Using the student's course grade as a measure of their knowledge, we look for differences in similarity of responses (to managers) across grades. Our assumption is that students who achieve an A or B (C) grade have more (less) knowledge than other students and therefore respond similarly (differently) to managers. Panel C of Table 5 shows that "A" students assessed a significantly lower likelihood that the impairment would be recorded (with a mean assessment of 4.18) than the managers. The difference is significant with a p-value of .06. This difference does not extend to students who received a "B" or "C" in the course suggesting that students earning higher grades may be poor surrogates for managers.

To explain why our results are not aligned with our expectations with respect to cognitive skills, we examine the mean age and work experience of these student groups (not tabulated). While A grade students were younger and had less work experience than B or C students, these differences are only significant when comparing the age of A to C students. We interpret this to mean the cognitive dimension captured by grades is incremental to our findings on work experience and possibly correlated with an age effect predominantly found in A students. Unfortunately, as ascertained in Table 2, the students earning A course grades are more likely to participate in experiments.

5.2.3 Culture

Cultural traits may lead to differential assessments of the likelihood that an asset impairment will be recorded therefore culture may be an important factor when determining which students to use as surrogates for managers. For the experiment employed in this paper, we expect North American managers' responses to be similar to responses from North American students, and

different from responses of students who spent a formative part of their lives outside North America. Prior literature suggests that the first fifteen years of a person's life are culturally the most formative (Parry & Urwin, 2011), therefore we use this metric to test whether North American students respond more similarly to the managers in Trottier (2013) than do foreign students.

Panel D of Table 5 provides the student sample responses divided into four categories based on where each student spent their first 15 years of life: North America, China, Other Asia, or Europe. Results show none of the sub-sample means are significantly different from the manager mean. Consequently, their country of origin[6] did not lead participants to perceive our financial reporting scenario differently, which suggests culture may not affect whether students are good surrogates for managers. We think this result could be due to the students all attending the same university and taking the same courses for a minimum of two years, with most students completing the same four-year program, therefore we suggest that researchers recruit students with several years of university experience in the geographical location of the managers they are meant to represent.

5.2.4 Gender

Research has found gender to be a factor affecting individuals' responses (e.g., Bernardi & Arnold, 1997; Lopez et al. 2005; Albaum & Peterson, 2006; Atakan et al. 2008; Eweje & Brunton, 2010). By partitioning the student and manager samples along gender lines, Panel E allows exploration of whether females were differentially skeptical. The panel provides statistics for each sample by gender. The male vs. female difference in response is tabulated in the last row. The manager vs. student difference is in the last column. Results show that male managers assessed a mean likelihood (5.24) that is significantly higher than female managers (3.55) with a difference of 1.69 that is significant at the 1% level. Similarly, male students provided a response

(4.87) that was significantly higher (p = .01) than their female cohorts' (3.84). However, male managers responded similarly to male students, with a difference in means of 0.37 (p-value = .26). Female manager and female student response differences are also insignificant (p-value = .55). We explore the possibility these results are caused by an uneven distribution of females among the four scenarios. For each of the two samples, we perform a Chi-square test to determine whether a disproportionate number of females were assigned scenarios with low likelihood impairments (such as bonus and/or permanent impairment), and find this is not the case.

Ultimately, in spite of gender differences, our inference is not overly distorted by utilizing students. However, we see reason to be cautious given that women accounted for a much larger proportion of the student subjects (approximately 51%) than manager subjects (approximately 22%), and that they provided significantly lower scores than men. It may be that the mismatched gender composition is what generated the overall difference in mean scores between students and managers. Hence, researchers should consider how gender composition could affect their inference when using students as surrogates for a target population, especially where males and females respond differently.

5.2.5 Conclusions on additional analysis

The analysis in Table 5 provides evidence that some sub-groups of students correspond better with managers. However, as we learned in our main findings, the blind application of statistical cutoffs is not the best approach to exploring the question of student surrogacy. Therefore, we suggest researchers consider our findings when constructing their surveys, and incorporate questions on work experience, student grades, age and gender to perform sensitivity analysis. A large and significant difference in responses along these dimensions indicates that results may not accurately project the target population responses.

5.3 Social desirability bias

In this study, we do not ask the respondents whether *they* will record the asset impairment. Instead the case scenario deals with how they think *another* person (Fred) will act. The viewpoint of this question has strengths and weaknesses. By asking what respondents think *another* person will do, we reduce the social desirability but induce some uncertainty as to what drives their responses. There could be many reasons why one participant expects Fred to record the loss while another does not. Intuitively, it could be that participants have different morals, philosophies, and ethics and project their choice of action onto Fred. Alternatively, it could be that they have different levels of distrust for Fred. The distrust would likely be based on Fred's gender or his position as manager, since there is no other information on Fred in the scenario. Mistrust based on position may very well be what drives the difference in responses between our student and manager samples, making the students more skeptical the impairment will be recorded. Mistrust based on gender may explain the differences in responses between male and females in the study, either because females are more skeptical of men or more skeptical in general.

In an ethics setting, other studies that explore gender and skepticism are scarce. A few papers study skepticism more generally. Examining male and female satisfaction with online shopping, Rodgers and Harris (2003) find men are generally more satisfied with their experience compared to women and attribute this difference to females being less trusting (more skeptical) of the situation. We leave it to future research to explore how these differences manifest when females participate in surveys.

6. Discussion

6.1 Summary

Research examining the appropriateness of students as surrogates is sparse and has produced

mixed results. We contribute to the literature by exploring the similarities of responses to an experiment conducted on managers and undergraduate students. The experiment is accounting in nature, as it requires participants to make an assessment of the likelihood that asset impairment will be recognized in the financial statements. A comparison of student and manager responses shows that, while the students appraised a significantly lower likelihood that the impairment would be recorded, the pattern of their responses was similar enough to the manager responses to lead to the same inference and conclusions. This suggests that when evaluating the use of students as surrogates for managers, researchers should take care to examine not only whether responses are significantly different, but whether these differences affect the conclusions.

6.2 Contribution to the literature

This paper adds to the literature by exploring the validity of drawing inferences from research experiments that use students as surrogates for managers. A second contribution comes from validating the original study by testing whether results hold with a different sample (Lindsay & Ehrenberg, 1993; Lindsay, 1995). Similar to Elliott et al. (2007), our study contributes by documenting attributes that make students better surrogates to managers.

6.3 Applied implications

Our findings have practical implications for researchers who recruit students for their experiments. Although students were appropriate surrogates in our setting, we find that students who have work experience provide responses that are more similar to management responses. Surprisingly, we find evidence that the best surrogates are students who obtain a B or a C grade, rather than the "A" students. Since a disproportionate number of "A" students participated in our survey, we conclude that grade composition within the student sample could bias the results of the experiment. A second potential source of bias is the gender composition. The gender composition of the student sample can be significantly different from the composition generally

found in the population of managers, which could lead to incorrect inference to the extent that women respond differently from men. Overall, our findings should provide researchers some guidance on the attributes of students to consider when selecting their convenience samples from the pool of undergraduate students. In particular, experimental researchers are encouraged to perform sensitivity analysis on their data to ensure that the gender and student grade composition of their sample does not bias their results.

6.4 Limitations and future research directions

Opportunities for future research arise from the limitations of others' studies. As with many research studies ours has limitations which may provide such opportunities. Our experiment examines one specific task related to financial reporting. Other experiments could use tasks such as budgeting, reporting or implementation of strategy through financial reporting to further explore use of undergraduate students as surrogates for managers. Additionally, we used students' course grade as a proxy for knowledge/intelligence, and obtain the unexpected result that "A" students are weaker surrogates. Perhaps a different measure such as overall GPAs (a measure not available to us due to protection of privacy laws in our jurisdiction) could be used to corroborate our finding. Our gender result establishes yet another path for future research: The source of the significant female skepticism could be further explored. Finally, our findings should be generalized with caution since a far greater number of experiments are needed for us to fully understand the benefits and limitations of using students in general and undergraduate students in particular as proxies. We encourage researchers who conduct experiments on managers to consider running a parallel experiment on students to advance the literature on the validity of students as surrogates.

Appendix . Research instrument.

A survey was created in four versions:

Version 1: permanent impairment with bonus plan

Version 2: reversible impairment with bonus plan

Version 3: permanent impairment with no bonus plan

Version 4: reversible impairment with no bonus plan

The existence of a bonus plan was incorporated in Versions 1 and 2 by inserting the sentence: **"**The corporation has a bonus plan, where the bonus is based on net income." in the space identified as **[1]** below. The reversible impairment aspect was incorporated in Versions 2 and 4 by inserting the sentence: "Suppose that recording an impairment loss is reversible; the loss can be reversed if the asset subsequently recovers its value, resulting in an increase in the carrying value of the asset on the balance sheet, and a corresponding increase in net income." in the space identified as **[2]** below. The permanent impairment aspect was incorporated in Versions 1 and 3 by inserting the sentence: "Suppose that recording an impairment loss is permanent; the loss cannot be reversed even if the value of the asset recovers in the future." in the space identified as **[2]** below.

The four versions of the research instrument were loaded onto a secure website and a unique online link obtained for each of them. Beta tests of the links confirmed that each instrument worked as designed and took 8-10 minutes to complete. Potential participants were identified, then randomized into four groups. Each group was sent an email link to one of the four surveys. Responses were received by an independent administrator. For the student sample, the administrator ran a program to match the data with original information while maintaining student anonymity. Hence each student's grade, course instructor, and course time (day/evening) were matched to his/her responses but no other information was kept.

Overall Instructions

Please read the following scenario and answer the questions as they appear on your screen. You should choose responses that reflect what you believe a manager would do in the situation described. Don't get too encumbered by the financial statement numbers; they are only there to provide context.

Scenario

Fred is the manager of a public corporation. He is responsible for preparing financial statements for reporting purposes but is not concerned about tax effects. [1] The corporation has a considerable investment in property, plant, and equipment as shown in last year's balance sheet:

Balance Sheet (2009)

ASSETS			LIABILITIES AND EQUITY	
Cash and Short-term Investments		$200,000	Accounts payable	$160,000
Accounts Receivable		$14,000		
Inventory		$350,000		
Property, Plant, and Equipment:				
Property		$2,000,000		
Plant and Equipment (Gross)	$5,000,000			
Accumulated Depreciation	($2,000,000)			
Net Plant and Equipment		$3,000,000	Equity	$5,404,000
Total Assets		**$5,564,000**	**Total Liabilities and Equity**	**$5,564,000**

Income Statement (2009)

Revenues	$885,000
Cost of Goods Sold	($80,000)
Depreciation	($100,000)
Operating Income	$705,000
Loss on Operations	($12,000)
Gain on Operations	$2,000
Net income	$695,000

It is the end of fiscal 2010 and Fred is preparing financial statements. Based on the frequent machinery breakdowns this year, he estimates that the net value of the plant and equipment (recorded at $3,000,000 in 2009) has declined beyond the amount that would be recorded after depreciation this year. His best estimate of the loss in value is approximately $300,000. However, the value may recover. A new method of retrofitting similar machinery is being developed, and if the retrofit becomes available it could bring the performance of the machinery to near new levels at a very low cost. Fred assesses a 5% probability that the value of the equipment will recover in the next three years due to retrofitting.

Accounting rules state that assets should be recorded at current value. Fred must decide whether to record an impairment loss, which would reduce the 'Net Plant and Equipment' on the balance sheet and reduce 'Net Income' through 'Loss on Operations'. [2]
In your opinion, how likely is Fred to record the impairment loss?

1	2	3	4	5	6	7
Not likely						Very likely

References

Abdel-Khalik, A.R. (1974). On the efficiency of subject surrogation in accounting research. *Accounting Review*, 49(4), 743-750.

Albaum, G., & Peterson, R.A. (2006). Ethical attitudes of future business leaders: Do they vary by gender and religiosity? *Business & Society*, 45(3), 300-321.

Alpert, B. (1967). Non-businessmen as surrogates for businessmen in behavioral experiments. *The Journal of Business*, 40(2), 203-207.

Ashton, R.H. (1974). An experimental study of internal control judgments. *Journal of Accounting Research*, 12(1), 143-157.

Ashton, R.H., & Kramer, S.S. (1980). Students as surrogates in behavioral accounting research: Some evidence. *Journal of Accounting Research*, 18(1), 1-15.

Atakan, M.G., Burnaz, S., & Topcu, Y.I. (2008). An empirical investigation of the ethical perceptions of future managers with a special emphasis on gender – Turkish case. *Journal of Business Ethics*, 82(3), 573-586.

Bean, D.F., & D'Aquila, J.M. (2003). Accounting students as surrogates for accounting professionals when studying ethical dilemmas: A cautionary note. *Teaching Business Ethics*, 7(3), 187-204.

Berger, K A., & Malinowski, C. (2004). Differences in ethical attitudes of undergraduates of Asian and US ancestry. *Advances in Competitiveness Research*, 12(1), 79-93.

Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticisms of laboratory experiments. *American Psychologist*, 37(3), 245-257.

Bernardi, R.A., & Arnold, D.F. (1997). An examination of moral development within public accounting by gender, staff level and firm. *Contemporary Accounting Research*, 14(4), 653-668.

Bernardi, R.A., Witek, M.B., & Melton, M.R. (2009). A four-country study of the associations between bribery and unethical actions. *Journal of Business Ethics*, 84(3), 389-403.

Bonner, S.E. (2008). *Judgment and Decision Making in Accounting*. Upper Saddle, NJ: Pearson Education Inc.

Brozovsky, J.A., & Sopariwala, P.R. (1995). The determinants of performance plan adoption. *Journal of Accounting, Auditing & Finance*, 10(3), 655-676.

Brunton, M., & Eweje, G. (2010). The influence of culture on ethical perception held by business students in a New Zealand university. *Business Ethics: A European Review*, 19(4), 349-362.

Byrnes, J., Miller, D., & Schafer, W. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, 125(3), 367-383.

Chang, C.J., & Ho, J.LY. (2004). Judgment and decision making in project continuation: A study of students as surrogates for experienced managers. *Abacus*, 40(1), 94-116.

Chung, J., & Monroe, G.S. (2003). Exploring social desirability index, *Journal of Business Ethics*, 44, 291-302.

Copeland, R.M., Francia, A.J., & Strawser, R.H. (1973). Students as subjects in behavioral business research. *The Accounting Review*, 48(2), 365-372.

Dickhaut, J., Livingstone, J., & Watson, D. (1972). On the use of surrogates in behavioural experimentation. *The Accounting Review*, 47(Supplement), 455–470.

Druckman, J.N., & Kam, C.D. (2011). Students as experimental participants: A defense of the 'Narrow Data Base.' In J.N. Druckman, D.P. Green, J.H. Kuklinski, & A. Lupia (Eds.), *Cambridge handbook of experimental political science* (pp. 41-57). New York: Cambridge University Press.

Duru, A., Mansi, S.A., & Reeb, D.M. (2005). Earnings-based bonus plans and the agency costs of debt. *Journal of Accounting & Public Policy*, 24(5), 431-447.

Elliott, W.B., Hodge, F.D., Kennedy, J. J., & Pronk, M. (2007). Are MBA students a good proxy for nonprofessional investors? *The Accounting Review*, 82(1), 139-168.

Enis, B.M., Cox, K.K., & Stafford, J.E. (1972). Students as subjects in consumer behavior experiments. *Journal of Marketing Research*, 9(1), 72-74.

Eweje, G., & Brunton, M. (2010). Ethical perceptions of business students in a New Zealand university: do gender, age and work experience matter? *Business Ethics: A European Review*, 19(1), 95-111.

Fehr, E., & List, J.A. (2004). The hidden costs and returns of incentives-trust and trustworthiness among CEOs. *Journal of the European Economic Association*, 2(5), 743-771.

Fisher, R.R. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, 20, 303-315.

Frederick, D. (1991). Auditors' representation and retrieval of internal control knowledge. *The Accounting Review*, 66(2), 240-258.

Fuchs, S., & Sarstedt, M. (2010). Is there a tacit acceptance of student samples in marketing and management research? *International Journal of Data Analysis Techniques and Strategies*, 2(1), 62-72.

Gift, M.J., Gift, P., & Zheng, Q., (2012). Cross-cultural perceptions of business ethics: Evidence from the United States and China. *Journal of Business Ethics*, 114(4), 633-642.

Globiel, J., & Phillips, F., (2001). Relating case presentation style and level of student knowledge to fact acquisition and application in accounting case analyses. *Issues in Accounting Education*, 16(2), 205-222.

Graham, J.R., Harvey, C.R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics*, 40(1), 3–73.

Greenberg, J. (1987). The college sophomore as guinea pig: Setting the record straight. *The Academy of Management Review*, 12(1), 157–159.

Halpren, D.F. (2012). *Sex Differences in Cognitive Abilities*, 4th Edition. New York, NY: Psychology Press.

Healy, P.M. (1985). The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics*, 7(1/2/3), 85-107.

Hoffman, V.B., Joe, J.R., & Moser, D.V. (2003). The effect of constrained processing on auditors' judgments. *Accounting, Organizations and Society*, 28 (7/8), 699-714.

Holthausen, R.W., Larcker, D.F., & Sloan, R.G. (1995). Annual bonus schemes and the manipulation of earnings. *Journal of Accounting and Economics*, 19(1), 29-74.

Houghton, K.A., & Hronsky, J.J.F. (1993). The sharing of meaning between accounting students and members of the accounting profession. *Accounting & Finance*, 33(2), 131-147.

Hughes, C.T., & Gibson, M.L. (1991). Students as surrogates for managers in a decision-making environment: An experimental study. *Journal of Management Information Systems*, 8(2), 153-66.

James, W.L., & Sonner, B.S. (2001). Just say no to traditional student samples. *Journal of Advertising Research*, 41(5), 63-72.

Jensen, M.C., & Meckling. W.H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305-360.

Kitayama, S., Park, H., Sevincer, A.T., Karasawa, M., & Uskul, A.K. (2009). A cultural task analysis of implicit independence: Comparing North America, Western Europe, and East Asia. *Journal of Personality and Social Psychology*, 97(2), 236–55.

Kohli, A.S., Peng, C., & Mittal, P. (2011). Predictors of student success in undergraduate business statistics. *Journal of Academy of Business and Economics*, 11(4), 32-42.

König, C., Steinmetz, H., Frese, M., Rauch, A., & Wang, Z-M. (2007). Scenario-based scales measuring cultural orientations of business owners. *Journal of Evolutionary Economics*, 17(2), 211–239.

Lambert, R.A., & Larcker, D.F. (1987). An analysis of the use of accounting and market measures of performance in executive compensation contracts. *Journal of Accounting Research*, 25(3), 85-129.

Libby, R., & Frederick, D.M. (1990). Experience and the ability to explain audit findings. *Journal of Accounting Research*, 28(2), 348-367,

Libby, R., Bloomfield, R., & Nelson, M. W. (2002). Experimental research in financial accounting. *Accounting, Organizations and Society*, 27, 775–810.

Lindsay, R.M. & Ehrenberg, A.S.C. (1993). The Design of Replicated Studies. *The American Statistician*, 47(3), 217-228.

Lindsay, R.M. (1995), Reconsidering the status of tests of significance: An alternative criterion of adequacy. *Accounting, Organizations and Society*, 20(1), 35-53.

Liyanarachchi, G.A., & Milne, M.J. (2005). Comparing the investment decisions of accounting practitioners and students: An empirical study on the adequacy of student surrogates. *Accounting Forum*, 29(2), 121-135.

Lopez, Y.P., Echner, P.L., & Olson-Buchanan, J.B. (2005). Shaping ethical perceptions: An empirical assessment of the influence of business education, culture, and demographic factors. *Journal of Business Ethics*, 60(4), 341-358.

McCabe, D.L., Dukerich, J.M., & Dutton, J. (1993). Values and moral dilemmas: A cross-cultural comparison. *Business Ethics Quarterly*, 3(2), 117-130.

Mortensen, T., Fisher, R. & Wines, G. (2012), Students as surrogates for practicing accountants: further evidence, *Accounting Forum*, Vol. 36, No. 4, pp. 251-265.

Nelson, M. (1993). The effects of error frequency knowledge and accounting knowledge on error diagnosis in analytical review. *The Accounting Review*, 68(4), 804-824.

Olsen, R.A., & Cox, C.M. (2001). The influence of gender on the perception and response to investment risk: The case of professional investors. *Journal of Psychology & Financial Markets*, 2(1), 29-36.

Osgood, C.E., Suci, G.J., & Tannenbaum, P.H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.

Parry, E., & Urwin, P. (2011). Generational differences in work values: A review of theory and evidence. International *Journal of Management Reviews*, 13(1), 79-96.

Peterson, R.A. (2001). On the use of college students in social science research: Insights from a second-order meta-analysis. *Journal of Consumer Research*, 28(3), 450-61.

Remus, W. (1986).Graduate students as surrogates for managers in experiments on business decision making. *Journal of Business Research*, 14(1), 19-25.

Rodgers, S., & Harris, M.A. (2003). Gender and e-commerce: An exploratory study. *Journal of Advertising Research*, 43(3), 322-329.

Sears, D.O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51(3), 515-30.

Shuptrine, F.K. (1975). On the validity of using students as subjects in consumer behavior investigations. *Journal of Business*, 48(3), 383-90.

Simon, A. (1979). Effects of selective sampling on a factor analysis. *Journal of General Psychology*, 101(2), 259-64.

Smith, C. W., Jr., & Watts, R.L. (1982). Incentive and tax effects of U.S. executive compensation plans. *Australian Journal of Management*, 7(2), 39-157.

Trottier, K. (2013). The effect of reversibility on a manager's decision to record asset impairments. *Accounting Perspectives,* 12(1), 1-22.

Tubbs, R. (1992). The effect of experience on the auditor's organization and amount of knowledge. *The Accounting Review*, 67(4), 783-801.

Vera-Munoz, S.C., Kinney, W.R., & Bonner, S.E. (2001). The effects of domain experience and task presentation format on accountants' information relevance assurance. *The Accounting Review*, 76(3), 405-429.

Watts, R.L., & Zimmerman, J.L. (1986). *Positive accounting theory*. Englewood Cliffs: Prentice-Hall, Inc.

Weber, R. (1980) Some characteristics of the free recall of computer controls by EDP auditors. *Journal of Accounting Research*, 18(1), 214-241.

Zhang, Y., Straub, C., & Kusyk, S. (2007). Making a life or making a living? Cross-cultural comparisons of business students' work and life values in Canada and France. *Journal of Cross Cultural Management: An International Journal*, 14(3), 174-195.

Table 1: Assessment of students as surrogates literature

Article Authors, Date	**Dimensions /Mechanism**	**Experimental Setting /Analysis/Technique**	**Characteristics Examined**	**Student Sample Description**	**Surrogacy Conclusion**
Alpert 1967	Knowledge Attitudes	Argument generation	Experience	Undergraduates (U/Gs): 77 HR; 38 engineering	Poor surrogates
Enis et al. 1972	Perspective Attitudes	Consumer behaviour experiment; biases (e.g., race)	Life experience	210 marketing U/Gs	Inconclusive
Copeland et al. 1973	Attitudes	Questionnaire: Financial reporting practices	Knowledge	160 accounting U/Gs	Poor surrogates
Shuptrine, 1975	Consumer attitudes	Questionnaire: Scaling use of products	Gender Life Experience	50 female education graduate students	Inconclusive
Abdel-Khalik 1974	Decision making	Assessed 40 lending cases	Knowledge	13 to 18 MBA students per case	Poor surrogates
Ashton and Kramer 1980	Decision making	Experiment: internal control judgments	Knowledge Experience	30 auditing U/Gs	Adequate surrogates
Weber 1980	Decision making; recall	Experiment: Knowledge content and structure	Experience Knowledge	96 auditing U/Gs	Inadequate
Remus 1986	Decision making	Production scheduling	Education Experience	13 MBA students – little experience	Good surrogates
Libby and Frederick, 1990	Decision making	Understanding financial statement errors	Experience	70 advanced auditing U/Gs	Inadequate
Hughes and Gibson 1991	Decision making	Questionnaire pre-/ post-training in supports system	Knowledge Ability to learn	MBA students (no sample size provided)	Inadequate
Frederick 1991	Recall	Internal control; Recall of knowledge	Knowledge; Memory	97 auditing U/Gs	Inadequate

Tubbs 1992	Knowledge	Experiment: Knowledge of audit errors	Experience	23 introductory auditing U/Gs	Inadequate
Houghton and Hronsky 1993	Recognition; Measuring meaning	Experiment: Recognition of accounting concepts; measuring those concepts' meaning	Knowledge Experience Age	132 advanced accounting U/Gs	Good surrogates in recognition; not meaning
Nelson, 1993	Understanding and knowledge	Financial statement errors and ability to structure knowledge	Experience Knowledge Learning	112 intermediate accounting U/Gs	Inadequate
James and Sonner 2001	Reactions Perceptions	Experiment: Advertisements	Age	116 full-time and 114 older part-time U/Gs	Good surrogates when similar age
Bean and D'Aquila 2003	Decision making	Experiment: Financial reporting cases with ethical considerations	Education; Experience	110 accounting U/Gs	Unsuitable
Hoffman et al. 2003	Decision making	Experiment: Going concern risk in two settings – constrained and unconstrained	Experience	35 accounting U/Gs in final week of auditing course	Good surrogates in constrained setting only
Chang and Ho 2004	Decision making	Experiment: Judgment regarding project continuation	Experience	146 business U/Gs	Unsuitable
Fehr and List 2004	Decision making	Experiment: Trust game	Experience	126 U/Gs	Unsuitable
Liyanarachchi and Milne 2005	Decision making	Experiment: Environmental liabilities disclosure investment	Education	51 final year accounting U/Gs	Good surrogates in this setting

Elliott et. al 2007	Decision making	Experiment: Investment decisions involving integrative complexity	Education Knowledge	82 early MBA; 42 advanced MBA students	Advanced MBAs better surrogates
Mortensen et al. 2012	Decision making	Experiment: Classification judgments	Education Experience	58 advanced acctg.; 60 engineering U/Gs	Accounting students good surrogates
Trottier and Gordon (present study)	Decision making	Experiment: Accounting impairment with ethical aspect	Experience Age Grades/ knowledge Gender	103 fourth year undergraduates who had completed an accounting theory course	Good surrogates except with respect to gender and grades

Table 2
Descriptive statistics

Panel A: Manager vs. Student Demographics *(number of observations)*[a]

	Managers	**Students**
Total	118	103
Gender		
Male	80	49
Female	22	51
Age		
Minimum	31	21
Mean	48	23
Maximum	65	40
Years of experience		
Minimum	9	0
Mean	26	4
Maximum	51	20

Panel B: Student Participation *(response rates in percentages)*

	Percentage	**Difference**	**p-value**
By time of class			
Day (n=82)	16.63%	3.81%	.25
Evening (n=21)	12.82%		
By grade[b]			
"A" (n=48)	27.75%		
"B" (n=43)	13.74%		
"C" (n=12)	6.94%		
"A" vs. "B"		14.01%	.00
"A" vs. "C"		20.81%	.00
"B" vs. "C"		6.80%	.06

[a] Some participants did not answer all demographics questions.
[b] Students with a grade lower than a "C" are not included in the analysis.

Table 3
Results of evaluation of the likelihood that loss impairments would be recorded.[a]

Panel A: Overall Mean and Standard Deviation

	Managers	**Students**	**p-value** [b]
Mean	4.85	4.37	*.07*
Standard Deviation	1.84	2.06	*.24*
	(n=118)	(n=103)	

Panel B: Mean by Treatment Combination

permanent impairment with bonus	3.93 (n=28)	3.31 (n=21)	*.08*
reversible impairment with bonus	5.25 (n=28)	5.07 (n=29)	*.50*
permanent impairment with no bonus	4.83 (n=24)	4.18 (n=32)	*.12*
reversible impairment with no bonus	5.24 (n=38)	4.75 (n=21)	*.30*

[a] The subjects evaluated the likelihood that the manager in the scenario would record the loss from 1 (very unlikely) to 7 (very likely).

[b] P-values relate to the difference between Manager and Student means and standard deviations. The number of observations (n) is in parentheses.

Table 4
ANCOVA estimates of reversibility and bonus effects.

Panel A

Estimated with manager and student in one sample together.

Model:[a] $Likelihood = \alpha + \beta_1 Reversible + \beta_2 NoBonus + \beta_3 Interaction + \delta_1 SDum + \delta_2 SDum*Reversible + \delta_3 SDum*NoBonus + \delta_4 SDum*Interaction + \varepsilon$

	Coefficient	Estimate	p-value
Intercept	α	3.94	*0.00*
Reversible	β_1	1.58	*0.00*
NoBonus	β_2	1.14	*0.03*
Interaction	β_3	-1.16	*0.09*
SDum	δ_1	-1.29	*0.02*
*SDum*Reversible*	δ_2	0.95	*0.07*
*SDum*NoBonus*	δ_3	0.18	*0.81*
*SDum*Interaction*	δ_4	-0.52	*0.61*
N		221	
adjusted R^2		20%	

Panel B

Estimated separately for manager and student samples.

Model:[b] $Likelihood = \alpha + \beta_1 Reversible + \beta_2 NoBonus + \beta_3 Interaction + \varepsilon$

		Manager Sample		Student Sample	
	Coefficient	Estimate	p-value	Estimate	p-value
Intercept	α	3.38	*0.00*	3.30	*0.00*
Reversible	β_1	2.05	*0.00*	1.76	*0.00*
NoBonus	β_2	1.63	*0.00*	1.17	*0.01*
Interaction	β_3	-1.65	*0.01*	-1.37	*0.04*
N		118		103	
adjusted R^2		19%		10%	

[a] This analysis corresponds to Equation 1 in the text. The variables are: Likelihood is the subject's assessment (from one to seven) of the likelihood the impairment would be recorded; the remaining variables are indicator variables equal to one if the respondent's scenario treated impairments as reversible (Reversible), if the manager in the scenario had no bonus plan (No Bonus), and if the plan was both reversible and without bonus (Interaction). The first two indicator variables are designed to have coefficients in the same direction to avoid having the interaction term capture the net result of a positive and negative effect. The remaining variables are a dummy variable if the participant was a student (SDum) and the interaction of this dummy with the first three variables in the model, to construct variables that measure the incremental student-effect on each treatment (SDum*Reversible, SDum*NoBonus, SDum*Interaction).

[b] This analysis corresponds to Equation 2 in the text.

Table 5
Comparison of manager and student responses

Comparison of manager mean response of 4.85 (to the likelihood the loss impairment would be recorded, on a scale of 1 to 7), to student responses based on student's work experience, course time, grade, country of origin, and gender.

Panel A: to Means of Students Based on Work Experience

	N	Mean	St.Dev.	Difference	p-value	Conclusion
0-4 years	63	4.03	2.42	.82	.02	different
5-9 years	31	4.98	1.87	-.13	.73	similar
10+ years	9	4.54	2.08	.31	.67	similar

Panel B: to Means of Students Grouped by Course Time

	N	Mean	St.Dev.	Difference	p-value	Conclusion
Day	82	4.24	2.10	.61	.04	different
Evening	21	4.87	1.95	-.02	.95	similar

Panel C: to Means of Students Grouped by Student Grade

	N	Mean	St.Dev.	Difference	p-value	Conclusion
"A"	48	4.18	2.10	.67	.06	different
"B"	43	4.38	2.08	.47	.20	similar
"C"	12	5.01	1.89	-.16	.78	similar

Panel D: to Means of Students Grouped by where they Spent First 15 Years of their Lives

	N	Mean	St.Dev.	Difference	p-value	Conclusion
North America	68	4.23	2.21	.62	.12	similar
China	20	4.41	2.11	.44	.36	similar
Other Asia	10	4.56	1.96	.29	.58	similar
Europe	4	5.93	1.26	-1.08	.28	similar

Panel E: Comparison of Responses by Gender[a]

	MANAGERS		STUDENTS			
	Mean	**St.dev.**	**Mean**	**St.dev.**	**Difference**	**p-value**
MALE	5.24 (n=80)	1.88	4.87 (n=49)	1.84	0.37	.26
FEMALE	3.55 (n=22)	1.82	3.84 (n=51)	2.10	-0.29	.55
Difference	1.69		1.03			
p-value	<.01		<.01			

[a] Some participants did not provide information on their gender, resulting in fewer observations available for analysis in Panel E of Table 5.

ENDNOTES

[1] We use the word "manager" generically here. In actuality, the "manager" sample consisted of participants in various positions who have significant influence on the content of their firms' financial statements, as described in Section 4.1 of this paper.

[2] The student survey contained 13 questions (in addition to the demographic questions) to be consistent with the manager survey. However, only the first question (regarding the likelihood Fred would record the impairment) is used in this paper.

[3] An ANOVA test (not tabulated) reported no significant difference in responses based on the number of terms remaining prior to a student graduating.

[4] The model in Trottier (2013) has some variables (such as firm revenue) that do not pertain to students. Her model requires ANCOVA to accommodate these continuous variables, while ours has only categorical variables and therefore could be tested with ANOVA. We use ANCOVA for comparability and interpretation.

[5] We noticed through detailed analysis that students who did not provide demographic information had significantly larger values of absolute differences between their likelihood score and the mean score of the managers within their treatment condition. There were not a sufficient number of these outliers to affect results, but it suggests there could be something systematically different about participants who fail to provide full information.

[6] We confirmed these results using the students' birth country (untabulated) to partition the students. While the means are different, the results remain the same.