

An Empirical Look At Developmental Interventions And Student Performance In The First Intermediate Accounting Course

Terry J. Ward, Middle Tennessee State University, USA
Rebecca A. Foote, Middle Tennessee State University, USA
Mary E. Phillips, Middle Tennessee State University, USA

ABSTRACT

Due to concerns about student performance in Intermediate Accounting I, our Department of Accounting established an Intermediate Readiness Committee in the spring of 2006 to create a developmental program for students entering Intermediate I, with the goal of improving performance in that course. Over the next two years, the Committee established the Developmental Program with two escalating interventions to improve performance in Intermediate I. These interventions were comprised of readiness testing with study sessions and readiness testing with use of developmental software. This study reports the impact these interventions had on student performance in Intermediate I. The authors control for gender, grades earned in each of the two accounting principles courses, whether the student took Principles II at school of study or transferred the course credit from another school, and overall undergraduate grade point average upon entering Intermediate I. Results indicate that each intervention resulted in significantly higher grades than observed in a Base Period without intervention.

Keywords: Intermediate Accounting I; Developmental Intervention; Readiness Test; Developmental Software

INTRODUCTION

Many students find passing the first intermediate accounting course a difficult task (Waples and Darayseh, 2005; Kilpatrick et al., 2011), and this course is often the deciding factor in whether a student remains an accounting major or changes to another discipline. We found students enter Intermediate I with varying levels of preparedness. In order to improve performance in Intermediate I, our Department of Accounting decided to develop a program to assess readiness and develop students for Intermediate I. The objective of this study is to determine the impact our Developmental Program interventions had on student performance in the first intermediate accounting course.

In this paper, we extend prior research by testing the impact of escalating interventions on subsequent student performance in Intermediate I. These interventions were comprised of readiness testing with study sessions and readiness testing with the use of developmental software. We were unable to find prior accounting education studies that used our approach to testing and developing accounting students entering Intermediate I.

Results suggest that our Developmental Program was successful in improving student performance in Intermediate I. All interventions resulted in significantly higher grades in Intermediate I when compared to a base period without intervention. The study also improves on the methodology of prior literature related to improving performance in Intermediate Accounting (Danko-McGhee and Duke, 1992; Huang et al., 2005; Shoulders and Hicks, 2008) by using additional control variables in multivariate models, compared to simpler models in previous studies (t-tests, Mann-Whitney difference of proportions, simple linear regression, etc.). The authors control for gender, grades earned in each of the two accounting principles courses, whether student took Principles II at school

of study or transferred in that course credit, and overall undergraduate grade point average upon entering Intermediate I.

This paper is organized as follows. First we present background information and an explanation of our program. Next we discuss prior research and present our hypotheses, followed by a description of our sample and methodology. Finally, we provide an analysis of our results and present our conclusions, recommendations and suggestions for future research.

BACKGROUND AND PROGRAM DESCRIPTION

Motivated by a desire to improve the success rate in Intermediate I, and concerned about the variability in students' knowledge coming from Principles II, we formed an Intermediate Readiness Committee in the spring of 2006. The Committee included faculty who were either teaching, or had recently taught, Intermediate I. Our concern for student variability of knowledge stems from the diversity of our student body. We are a large public institution with a substantial number of transfer students coming from community colleges and other universities. Also, we have a large population of adult students returning to school, some of whom have a lag of several semesters, or years, since taking Principles II.

Total enrollment at the institution where this study was conducted was 23,872 students in the fall of 2008, with undergraduates comprising 89 percent of total enrollment. Seventy-seven percent of the students were enrolled full time, women were a majority at 54 percent, and minority students represented 22 percent of the University's enrollment. The average age of the undergraduate student population was 23 years and the average ACT score was 22.1, while the national average was 21.1 for that year.

Interventions

The Intermediate Readiness Committee implemented a Readiness Test as a course requirement beginning fall of 2006. A letter that explained the Readiness Test Policy, and advised students that the test would be administered during the first class meeting, was sent to students prior to the start of the semester. Passing the Readiness Test required a score of 70 percent or higher. Failing the test required retesting. Students were permitted to retake the test up to four times prior to an announced deadline, which was five and a half weeks into the semester. After each test, the faculty gave non-mandatory review sessions for students needing to retake the test. Multiple testing opportunities with review sessions represented the intervention process for that semester.

The Readiness Test consisted of 40 multiple choice questions covering financial accounting content from Principles I and II, with emphasis on the accounting cycle. The first four tests were unique, while the last one was drawn from questions in the prior tests. Students receiving less than 70 percent on the initial test were not required to retest at each of the four additional opportunities until passing, but they were required to obtain a passing score by the final retake opportunity.

In addition to the review sessions that were provided between the retake opportunities, the Intermediate Readiness Committee also provided individualized feedback after the second retake opportunity to all students who had failed the test at least two times, and allowed students to review their prior tests. After the fourth retake, nine students (7.75%) failed to pass the test and were forced to withdraw from the class. This period is labeled the Readiness Test and Review Sessions Period of our study.

The Intermediate Readiness Committee analyzed the results of the fall 2006 semester and decided to replace the review sessions component of the process with developmental software and limit the testing to one test of readiness. The multiple retake and review processes had been frustrating and time consuming for students. Time spent attempting to pass the test took away needed study time for Intermediate I topics.

To address these concerns, beginning spring of 2007, the Intermediate Readiness Committee gave the Readiness Test to determine the level of student knowledge prior to entering Intermediate I. Students were permitted one attempt at the test. Those who passed the test were not required to complete any additional

developmental work. However, students who failed the Readiness Test were required to complete a developmental software learning module.

During all subsequent semesters, a study guide was posted on the departmental website, and testing opportunities for the Readiness Test were available well in advance of the beginning of each semester. Students received test scores within 24 hours, so they knew quickly whether they had passed the Readiness Test. In the event a passing score was not obtained, students were required to complete a developmental software learning module.

After seeing a demonstration at the American Accounting Association annual meeting, we selected ALEKS[®] Accounting Cycle, an online learning module, to serve as our developmental software component. Students were required to purchase the software the first two semesters that we utilized ALEKS[®], but we were able to obtain funding so there was no individual cost to students in subsequent semesters. Students who did not pass the Readiness Test were required to complete ALEKS[®] through the Goal Completion Assessment, which is the final element of the module.

The ALEKS[®] Accounting Cycle module contains 54 topics, 40 of which address the accounting cycle and an additional 14 that address merchandising and inventory. Our definition of successful completion of ALEKS[®] requires a student to work through every one of the 54 topics included in the module, followed by completion of the Goal Completion Assessment to the best of the student's ability. Each of the 54 topics is represented by a corresponding piece of pie in a pie chart that the student must completely fill before the Goal Completion Assessment will trigger. Many of the topics are prerequisites of other topics. For each of the 54 topics, students are presented with a task, such as completing a balance sheet or calculating gross profit. If a student believes that he or she knows how to complete the task, the student may do so, and then ALEKS[®] grades the student's work. The student is not awarded the piece of pie that corresponds with the task until the student successfully completes the respective task, and then demonstrates proficiency two more times when presented with algorithmic versions of the original task. Once the student has demonstrated proficiency of the task three separate times, the corresponding piece of pie is added to the student's pie chart.

The Goal Completion Assessment does not trigger until the student has worked through all 54 topics, proving proficiency on each topic under three different sets of facts. According to our data, students spent, on average, 15 hours to complete the pie chart and the Goal Completion Assessment. Also, the average score on ALEKS[®] was 91 percent, with a standard deviation of 15.65.

The deadline for completion was approximately two and a half weeks into the semester, which coincided with the last day to drop a course without a grade. If a student failed to complete ALEKS[®] by the deadline then he/she had to either drop the course or receive a failing grade for the semester. This period is labeled the Readiness Test and Developmental Software Period of this study.

We continued utilize the Readiness Test to identify weaker students, coupled with the use of developmental software. The Readiness Test is basically the same across all the periods of intervention, with only the mix of questions that comprise each exam changing. The following variables were held constant across all periods of the Readiness Test:

1. Multiple choice format;
2. All questions come from a test bank developed specifically for Readiness Test purposes by the Intermediate Readiness Committee;
3. Tests financial accounting topics from traditional Principles of Accounting courses;
4. Score of 70 percent required to pass;
5. Responsibility for exam security resides with Program Coordinator, who is the only person who copies, grades and stores the exams;
6. Allotment of one hour for the student to complete the exam; and
7. Scantron and calculator provided to the student.

PRIOR RESEARCH AND HYPOTHESES DEVELOPMENT

Prior research indicates that diagnostic exams are useful in predicting Intermediate I grades (McCormick and Montgomery, 1974; Buehlmann, 1975; Delaney et al., 1979; Hicks and Richardson, 1984; Danko-McGhee and Duke, 1992; Lynn and Robinson-Backmon, 2006). Shoulders and Hicks (2008) employed a repetitive diagnostic exam, where the instructor explained weaknesses at each iteration until a score of 80 percent or better is achieved, at which time the student was allowed to take the next interim exam in Intermediate I. Their results showed better performance and fewer withdrawals for Intermediate I. Waples and Darayseh (2005) found that a diagnostic exam score was, along with cumulative GPA, positively related to performance in Intermediate I. However, their diagnostic exam was given several weeks into the semester after some of the students had dropped the course. They recommended a diagnostic exam at the beginning of the semester, and also suggested having a required minimum GPA of 2.5 in the introductory principles courses.

Danko-McGhee and Duke (1992) found both overall GPA and a diagnostic exam significantly related to the Intermediate I grade. Turner et al. (1997) used the first exam in Intermediate I, along with GPA, major, and where the Principles courses were taken, to predict grades in Intermediate I. The above studies tended to use a diagnostic test as an assessment instrument to determine if a student had the skills necessary to succeed in Intermediate I. There was little attempt to identify students needing additional developmental help or to improve student performance through developmental interventions. Thus, the emphasis was on predicting success in Intermediate I, not improving performance in Intermediate I.

Overall GPA is a common explanatory factor in research related to success in Intermediate I (i.e. Frakes, 1977; Hicks and Richardson, 1984; Turner et al., 1997; Eikner and Montondon, 2001; Waples and Darayseh, 2005; Lynn and Robinson-Backmon, 2006; Burnett et al., 2010; Kilpatrick et al., 2011). However, prior research reports mixed results when investigating the relationship between GPA in the introductory accounting courses and grades in Intermediate I. Several authors found a positive relation between grades in Principles I and grades in Intermediate I (Frakes, 1977; Eckel and Johnson, 1983; Hicks and Richardson, 1984; Danko-McGhee and Duke, 1992; Kilpatrick et al., 2011; and Carrington, 2012). Others found that grades in Principles II are useful in predicting grades in Intermediate I (Buehlmann 1975; Delaney et al. 1979; Eckel and Johnson 1983). Waples and Darayseh (2005) found no relationship between grades in both Principles I or Principles II and grades in Intermediate I. Buehlmann (1975) found that grades in Principles of Accounting I are not significantly related to grades in Intermediate I. He credits this lack of relationship to the “accounting maturity” that is not reached by students until they complete Principles II (Danko-McGhee and Duke, 1992. p.2). Sanders and Willis (2009) considered other drawbacks to relying on Principles I grades as a predictor of success in Intermediate I. They suggested that students may have taken the course at another school with a less rigorous program, and there could have been a time lapse since taking the course.

Consistent with Sanders and Willis, other research indicates that it is important where the principles of accounting courses are taken, especially if taken at a four-year versus a two-year school. Several authors (Burdick and Schwartz, 1982; Turner et al., 1997; Kilpatrick et al., 2011) found that taking the introductory accounting courses at a four-year school is a predictor of success in Intermediate I. Other researchers found that students who took the principles courses at community colleges have significantly lower grades in Intermediate courses (Colley et al., 1996; Lynn and Robinson-Backmon, 2006) and upper division accounting courses (Laband et al., 1997; Carlan and Byxbe, 2000). Carlan and Byxbe (2000. p. 5) attributed the weaker performance of these two-year college transfer students to a combination of grade inflation and a “less nurturing and more demanding senior college environment.” However, transfer students did not appear to believe that where they took their principles courses impacted their performance in Intermediate I (Carrington, 2012). Carrington surveyed students regarding the elements that explained their success in Intermediate I; they did not perceive as significant the type of school where they took their accounting principles courses.

Prior accounting education studies tested for gender effects, although with mixed results. Several studies found that female students outperformed males in accounting exams (Mutchler et al., 1989; Lipe, 1989; Tyson, 1989), while other research found no significant gender differences (Doran et al., 1991; Gist et al., 1996; Turner et al., 1997; Eikner and Montondon, 2001; Davidson, 2002). Fogarty and Goldwater (2010) suggested that gender

study results are confounded when professors assign points to homework and other effort-based work, since their results indicate that females expend more effort.

Huang et al. (2005) performed screening testing or completion of a one hour remedial program to supplement a user perspective being taught in the Principles I course. They found that students who either passed a pretest or remedial course receive higher grades in Intermediate I.

Sanders and Willis (2009) used a Principles of Accounting Competency Exam (PACE) as a form of assessment. Sanders and Willis compared the semester prior to incorporating their program with the semester in which they incorporated the program, using the same three instructors in both semesters. Four remedial learning options were made available to students to help them review accounting topics; ALEKS[®] was one of the four options. Other learning options included self-study, a boot camp experience, and an opportunity to audit the first principles of accounting course. ALEKS[®] was not required and not used by all of the students. The assessment exam was the last step in the program, with two of three Intermediate I Professors using the exam as a required component of the course grade. The third Professor did not require students to take the assessment exam. They found that their program improved Intermediate I grades and reduced the number of withdrawals.

The authors of this study used the Readiness Test in a manner different than Sanders and Willis (2009), with major emphasis on improving performance in Intermediate I. We used the Readiness Test to first determine whether students possessed sufficient accounting knowledge to succeed in Intermediate I. Students who failed the Readiness Test were deemed to have insufficient knowledge to succeed in Intermediate I. These students were required to complete additional developmental work to continue in Intermediate I. In the first semester of intervention, after each test, the faculty gave optional review sessions for students needing to retake the test. For the other periods of intervention, students who did not demonstrate sufficient accounting knowledge, as measured by a minimum score of 70 percent on the Readiness Test, were required to complete ALEKS[®]. Completion of ALEKS[®] was required by the time the class completed the review chapters in the Intermediate I textbook (first four chapters). The intended goal of our program was to improve performance in Intermediate I, and the Readiness Test was used as an initial assessment of knowledge at the beginning of the process to determine which students needed additional developmental early in the Intermediate I.

Phillips et al. (2013) describe an Intermediate Readiness Program using assessment tests and ALEKS[®] and students' feedback regarding the program. Students believed that an assessment exam was useful and that the developmental program was needed. The students also liked ALEKS[®] and believed that the software helped them prepare for Intermediate I. The authors did not test the impact of the interventions on student performance, but only described their developmental program and student reactions to the interventions in order to help other schools develop similar intervention programs.

The purpose of this paper is to examine the impact on subsequent student performance in Intermediate I of various developmental interventions comprised of a Readiness Test with Review Sessions and Readiness Test with use of developmental software. Thus, this paper extends prior research in student performance in Intermediate I by analyzing various forms of interventions through the use of multivariate statistical models, by controlling for additional variables found significant in prior studies, and by including multiple semesters in our sample. The authors control for gender, grade earned in Principles I, grade earned in Principles II, whether the student either took the principles II course at the school of study or transferred the course from another school, and overall undergraduate grade point average upon entering Intermediate I.

To test whether our interventions improved performance in Intermediate I, we test the following hypotheses (in alternative form) in this study:

H1: The Readiness Test with Review Sessions Period results in significantly higher Intermediate I grades than in the Base Period;

H2: The Readiness Test with Developmental Software Period results in significantly higher Intermediate I grades than in the Base Period.

SAMPLE

The authors identified students who took Intermediate I from the fall of 2005 through the spring of 2008. We select the initial year of study for two reasons. First, this period coincided with the beginning of our review period for our Strategic Planning Committee, which led to the creation of the Intermediate Readiness Committee. Second, we believe this period of data collection allows us a sufficient baseline number of student observations to adequately compare before and after results of our interventions.

We segregate the data into three categories. Period 1, Base Period (no intervention) is the period in which no developmental plan was used, and includes data for fall of 2005, spring of 2006, and summer of 2006. Period 2, Readiness Test with Review Sessions Period, is the period in which the Department required students to earn a score of 70 percent or higher on a Readiness Test, with optional review sessions by faculty between tests, to remain in Intermediate I. Period 2 included data for fall of 2006. Period 3, Readiness Test with Use of Developmental Software Period is the period in which the department used the Readiness Test to first identify the students with weaknesses in accounting knowledge (those scoring less than 70 percent on one attempt on the test). Students failing to score 70 percent were then required to complete developmental software (ALEKS®) to remain in Intermediate I. Period 3 includes data for spring of 2007, summer of 2007, fall of 2007, and spring of 2008.

Our initial sample of students enrolled in Intermediate I during the periods resulted in 623 original observations, with 214 from Period 1, 107 from Period 2, and 302 from Period 3. These partitioned samples of student observations were used in all descriptive statistics and frequency counts when possible.

For the multivariate models developed in this study, 90 observations were dropped (36 from Period 1, 16 from Period 2, and 38 from Period 3). These students did not have sufficient data to complete all the variables used in the multivariate models. Thus, the final sample of student observations used in this study’s multivariate models contained 533 observations, with 178 from Period 1, 91 from Period 2, and 264 from Period 3. Although not reported in a table, the breakdown of male to female ratios remained almost equal across the samples.

METHODS

Response (Dependent) Variable

The response, or dependent variable, in this study is student performance in Intermediate I. Thus, the semester grade the student earned in Intermediate I during the period of study is the dependent variable in this study. This variable is coded as follows:

INTGRADE =	0	if student received a semester grade of F or W without another attempt in the period of study,
	1	if student received a semester grade of D,
	2	if student received a semester grade of C,
	3	if student received a semester grade of B, and
	4	if student received a semester grade of A.

Independent (Predictor) Variables

We create the following categorical variable measuring the effects of our interventions on performance in Intermediate I:

PERIOD	=	1	if student took Intermediate I during Period 1, Base Period;
	=	2	if student took Intermediate I during Period 2, Readiness Test with Review Sessions Period; and
	=	3	if student took Intermediate I during Period 3, Readiness Test with Developmental Software Period.

PERIOD is the variable of interest in this study. A significant positive result for PERIOD would indicate that one or more of our interventions in Intermediate I was successful in improving student performance. Contrasting the means between categories provides additional evidence of which interventions were successful.

As a result of prior research (e.g., Buehlmann, 1975; Delaney et al., 1979; Eckel and Johnson, 1983; Frakes, 1977; Hicks and Richardson, 1984; Lipe, 1989; Tyson, 1989; Davidson, 2002), and the type of university where this study takes place (large university with a diverse group of students), we selected the following control variables to include in each model:

- GENDER = 0 if student is female, and
1 if student is male;
- PRIN1_GRADE = grade student earned in Principles I;
- PRIN2_GRADE = grade student earned in Principles II;
- PRIN_OTHER = 0 if student took Principles II at the school where this study is conducted, and
1 if student transferred in the course grade for Principles II from another school; and
- TOTALGPA = overall undergraduate grade point average of all courses completed before entering Intermediate I.

GENDER is a dichotomous control variable to capture gender differences in performance. PRIN_OTHER is a dichotomous control variable to capture whether the student took Principles II at a school other than the one where this study is conducted. PRIN2_GRADE and TOTALGPA are continuous control variables designed to capture prior performance of the student.

The same professors taught Intermediate I across the four periods in this study. Each professor used the same textbook and covered the same course content. Professors used similar exams and a similar number of exams each semester over the period of our study, although there were no common exams administered. The authors also tested a control variable for each of the four professors who taught Intermediate I during the sampling period. Neither of the professor variables was significant, so these variables were excluded from the models reported in this study.

ANCOVA (ANOVA with Covariates) Model

For the initial multivariable model in this study, the authors treated the response variable for student performance, INTGRADE, as a continuous variable and created the following Analysis of Covariance (ANCOVA) model: $INTGRADE = PERIOD + GENDER + PRIN1_GRADE + PRIN2_GRADE + PRIN_OTHER + TOTALGPA$.

Since the model contains both categorical and continuous explanatory (independent) variables, it is an Analysis of Covariance (ANCOVA) model. ANCOVA models basically combine simple linear regression and Analysis of Variance (ANOVA) into the same model and are used when both nominal and continuous independent or explanatory variables are employed (as in this study). The Proc GLM function in SAS (2009) was used to create the ANCOVA model. Since students repeating Intermediate I can lessen the validity of results, we ran all models with these observations deleted from the sample.¹

The authors decided to treat INTGRADE as a continuous variable and use ANCOVA to develop the models for a couple of reasons. First, results from ANCOVA models are easier to explain when the explanatory variable of interest has multiple categories, as is the case for PERIOD (three categories). ANCOVA allows one to contrast mean differences and produce one statistic and significance for each variable. Categorical models based on

frequency probabilities are more difficult to interpret and more difficult to explain. Second, the use of ANCOVA generated models with grade treated as a continuous variable is more consistent with prior education research. Prior accounting education researchers have primarily used mean based models such as ANOVAs or ANCOVAs in their studies. The authors are not aware of empirical accounting education papers reporting empirical results from categorical models.

Since ANCOVA is based on analysis of means, it has similar assumptions as in linear regression. ANCOVA results are impacted by whether or not the data are normally distributed, the cells have approximately equal variances, and explanatory variables are impacted by auto-correlation. For all model results, the authors tested for normal distribution, impact of unequal variances in cells, and for auto-correlation. For the results reported in this study, the data meet the relevant assumptions and results were not negatively impacted.²

RESULTS

Grades by Period

Table 1 contains the frequency counts by letter grades that students earned in Intermediate I for each of the periods of this study, along with the Pearson chi-square statistic, which tests whether the proportions are significantly different. The frequencies suggest that the relationship between grades and the periods of study are in the direction expected and somewhat significantly different (Pearson chi-square of 16.72 with p-value of .08). A larger percentage of the students received As in each of the intervention periods than in the Base Period (7.5 percent versus 13.1, and 17.2 percent, respectively), while fewer students withdrew from the course. A positive pattern for the percentage of As and a negative pattern for the percentage of Ws appears across the periods of intervention, suggesting that each of the added interventions improved performance somewhat.

Table 1: Counts of Grades by Period of Observation

	Periods of Study			Totals
	Period 1: Base (No Intervention)	Period 2: Readiness Test and Review Sessions	Period 3: Readiness Test and ALEKS®	
Grade Earned in Intermediate I:	Count (%)	Count (%)	Count (%)	Count (%)
Withdrew and never earned a grade during period	24 (11.2)	12 (11.2)	28 (9.3)	64 (10.3)
F	32 (15.0)	13 (12.1)	31 (10.3)	76 (12.2)
D	17 (7.9)	6 (5.6)	24 (7.9)	47 (7.5)
C	77 (36.0)	31 (29.0)	85 (28.1)	193 (31.0)
B	48 (22.4)	31 (29.0)	82 (27.2)	161 (25.8)
A	<u>16 (7.5)</u>	<u>14 (13.1)</u>	<u>52 (17.2)</u>	<u>82 (13.2)</u>
	<u>214 (100.0)</u>	<u>107 (100.0)</u>	<u>302 (100.0)</u>	<u>623 (100.0)</u>

¹Pearson $\chi^2 = 16.72$ (p-value = .08, with 10 df)

²⁴

¹Pearson $\chi^2 = \sum_{i=1}^24 \chi^2_{ii}$, where $\chi^2_{ii} = [(y_{ii} - m_{ii})^2 / m_{ii}]$. y_{ii} is the observed count. m_{ii} is the expected or estimated count, calculated as $m_{ii} = [(y_{+i})(y_{i+})] \div n$.

Table 1 results suggest that the interventions primarily benefited students earning a grade of C or better, with smaller changes in the percentage of students earning failing grades. This is consistent with the findings of Sanders and Willis (2009). However, this observation is somewhat muddled by presenting the three failing grades (W, D, and F) separately. To better identify the trends in the grades, the authors collapsed the grades into three categories across the periods of study. The three categories are: (1) student received a W, D, or F; (2) student received a C; and (3) student received a B or A. This classification is logical given that students receiving a W, D, or F must repeat Intermediate I. Also, reducing the contingency table into three grade categories by period matrix

gives a statistical test of whether the downward trend in the C or lower grades and the upward trend in the B and A grades are significant. The result for this three grade category by period contingency matrix is shown in Table 2.

Table 2: Counts of Grades by Period of Observation using Three Grade Categories: Significance of Trends

	Periods of Study			Totals
	Period 1: Base (No Intervention)	Period 2: Readiness Test with Review Sessions	Period 3 Readiness Test with Developmental Software	
Grade Earned in Intermediate I:	Count (%)	Count (%)	Count (%)	Count (%)
W, F or D	73 (34.1)	31 (29.0)	83 (27.5)	187 (30.0)
C	77 (35.9)	31 (29.0)	85 (28.1)	193 (31.0)
B or A	<u>64 (30.0)</u>	<u>45 (42.0)</u>	<u>134 (44.4)</u>	<u>243 (39.0)</u>
	<u>214 (100.0)</u>	<u>107 (100.0)</u>	<u>302 (100.0)</u>	<u>623 (100.0)</u>

Pearson $\chi^2 = 11.555$ (p-value = .021, with 4df)

The frequencies reported for the collapsed grade variable in Table 2 illustrate a clearer picture of the impact of the interventions on the grades in Intermediate I, and trends in the percentages of grades received. The interventions appear to have had a significant escalating positive impact on the grade distributions. The interventions seem to improve the percentages of students receiving a B or an A, while decreasing the percentages of students receiving lesser grades (Pearson chi-square of 11.555 with p-value of .021).

The W, F, or D percentages declined each period, as did the percentages of the students receiving a C. The trend reverses for students in the B or A category. The interventions improved the percentages of students receiving a B or an A with each intervention. This result suggests that both of the interventions accomplished the objective of improving performance in Intermediate I, with the last intervention resulting in the greatest improvement over the Base Period.

Multivariate ANCOVA Models

The univariate results in Tables 1 and 2 suggest that each of the interventions produced some benefits for our students. However, results may differ after controlling for other information that could impact student performance. To better determine whether our interventions had incremental benefits, we developed multivariate models with numerous control variables.

To empirically test the hypotheses of this paper, the authors generated an Analysis of Covariance (ANCOVA) model by regressing the variable of interest representing the three periods of our study (PERIOD) and the five control variables (GENDER, PRIN1_GRADE, PRIN2_GRADE, PRIN_OTHER, TOTALGPA) on the measure of performance in Intermediate I (INTGRADE). Insignificant variables (p-value > 0.05) were then dropped, and another model with the significant variables was run again. Both full and reduced models’ results are reported in Table 3.

Table 3: ANCOVA Model
The Relationship between Periods of Readiness Intervention and Student Performance in Intermediate I

¹ Explanatory Variables	Sign of Relationship	Full Model			Reduced Model		
		DF ²	F statistic ³	P-value ⁴	DF	F statistic	P-value
PERIOD	+	2	5.61	0.003	2	5.71	0.003
GENDER	-	1	0.57	0.452			
PRIN1_GRADE	+	1	0.96	0.326			
PRIN2_GRADE	+	1	26.68	0.000	1	34.77	0.000
PRIN_OTHER	-	1	8.55	0.003	1	7.75	0.005
TOTALGPA	+	1	26.84	0.000	1	41.91	0.000
Overall Model F statistic		7	29.22	0.000	5	40.87	0.000

¹ PERIOD is the variable of interest in this study and was coded: 1 if student took Intermediate I during Period 1, Base Period; 2 if student took Intermediate I during Period 2, Readiness Test and Review Sessions Period; and 3 if student took Intermediate I during Period 3, Readiness Test and Developmental Software Period. GENDER is a dichotomous control variable to capture gender differences in the model, coded: 0 = female, and 1 = male. PRIN1_GRADE = average grade student earned in Principles I. PRIN2_GRADE = average grade student earned in Principles II. PRIN_OTHER is a dichotomous control variable to capture whether or not the student took Principles II at a school other than the one where this study is conducted, coded: 0 = student took Principles II at school where this study is conducted, and 1 = student took Principles II at another school. TOTALGPA = overall undergraduate grade point average of all courses completed before entering Intermediate I.

²DF are the degrees of freedom.

³F statistics tests how well the overall model and individual variables account for the response variable's behavior and is based on Type III, or partial sums of squares.

⁴P-value is the probability of arriving at the F statistic by chance occurrence.

The full ANCOVA model results show that GENDER and PRIN1_GRADE were not significant in explaining the grade earned in Intermediate I by the students (p-value not $\leq .05$). The relationship between PRIN1_GRADE and INTGRADE was in the direction expected; however, it was not significant (p-value of 0.326) when controlling for the other variables, and is consistent with earlier research by Buehlmann (1975). This result suggests that the grade a student earns in Principles I is not a good incremental predictor of how well the student will perform in Intermediate I.

GENDER is not significant, indicating that males and females performed similarly in Intermediate I (F statistic of 0.57, with p-value of 0.452). This finding is consistent with Doran et al. (1991), Gist et al. (1996), and Davidson (2002), but contrary to results reported by Mutchler et al. (1989), Lipe (1989), and Tyson (1989).

To save degrees of freedom, the insignificant variables (p-value > 0.05) were dropped and the model was run once more resulting in a reduced model with four explanatory variables. Results for the reduced model show that all of the remaining variables were significant and in the direction expected in explaining student grades in Intermediate I.

The significant result for PRIN_OTHER indicates that students who took Principles II at other schools and transferred in the credit for that course earned significantly lower grades than those who took that course at the school of this study.

Total grade point average on prior work (TOTALGPA) was the strongest variable in explaining INTGRADE (F statistic of 41.91, p-value = 0.000) and was positively related to INTGRADE; the higher the undergraduate grade point average of the student, the higher the grade the student earned in Intermediate I. As expected, the significant result for PRIN2_GRADE (F statistic of 34.77, with p-value = 0.000) indicates students with higher grades in Principles II earned significantly higher grades in Intermediate I. Results for PRIN2_GRADE and TOTALGPA are consistent with previous research and were expected (Buehlmann, 1975; Delaney et al., 1979; Eckel and Johnson, 1983).

The main goal of this study was to determine whether our interventions improved student performance in Intermediate I. We created a categorical variable, PERIOD, to capture the differences in interventions across our sample periods. The results for PERIOD indicate the periods did matter in explaining Intermediate I grades, even after controlling for other information.

The relationship between PERIOD and INTGRADE is positive and the test statistic is highly significant (F statistic of 5.71, p-value = 0.003), indicating that the average grades received by the students significantly differed across the periods of this study, and in the direction expected. Students tended to earn higher grades in the intervention periods than in the Base Period.

Analysis of the categories within the PERIOD variable is needed to test the hypotheses. To better examine these differences between categories of PERIOD, the authors developed a base Analysis of Variance model with PERIOD regressed on INTGRADE and created contrasts between the cells of PERIOD to determine which intervention(s) improved Intermediate I performance over the Base Period grades. The INTGRADE means by period of study with Analysis of Variance contrasts are reported in Table 4.

Table 4: Means and Contrasts with Statistics for Intermediate I Grade (INTGRADE) by Each Period of Study (PERIOD)

I. Panel A: Mean of Student Grades in Intermediate I by Period of Study:						
	Period 1: Base (No Intervention)		Period 2: Readiness Test with Review Sessions		Period 3: Readiness Test with Developmental Software	
¹ Variables	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
INTGRADE	1.752	1.313	2.094	1.368	2.208	1.378

II. Panel B: Contrasts Differences in Performance by Period of Study:		
	F Value	P Value ¹
Period 1 vs Period 2, test of 1.752 vs 2.094	3.87	0.049
Period 1 vs Period 3, test of 1.752 vs 2.208	11.43	0.000
Period 2 vs Period 3, test of 2.094 vs 2.208	0.49	0.485

¹Contrasts generated from simple ANOVA model with Period regressed on INTGRADE. One-tail p-value for direction testing on all contrasts.

Table 4 results show that each intervention resulted in improved average grades in Intermediate I (INTGRADE means of 1.752 for the Base Period, 2.094 for Readiness Test with Review Sessions Period, and 2.208 for Readiness Test with Developmental Software Period). To determine whether the interventions improved student performance in Intermediate I, the contrasts between the average Intermediate I grades for the Base Period and the two periods of intervention were developed. These are shown in Panel B of Table 4.

The contrasts suggest that both periods of intervention resulted in significantly higher grades than in the Base Period (p-value ≤ .05). The Readiness Test with Review Sessions intervention improved grades in Intermediate I over the grades in the Base Period (F value of 3.87, p-value of 0.049). When the Readiness Test was used to identify weak students and developmental software was then required (Period 3), Intermediate I grades were significantly greater than observed in the Base Period (F value of 11.43, p-value = 0.000). Thus, both H₁ and H₂ are accepted. Both forms of intervention improved Intermediate I grades significantly over those for the base period of observation.

Although not tested for hypotheses purposes, Panel B also contains a comparison of the two intervention periods. The grade means for each period indicate that the second intervention with the developmental software increased student performance (means of 2.094 versus 2.208, respectively), but the improvement was not significant at p-values ≤ .05. This comparison of the interventions with each other suggest that either form of intervention

would be beneficial as developmental programs for accounting departments at other universities wishing to improve student performance in Intermediate I.

CONCLUSIONS

Due to concerns about student performance in Intermediate Accounting I, our Department of Accounting created an Intermediate Readiness Committee with the goal of improving performance in Intermediate I. Over the next two years, the Committee established a Developmental Program with two escalating interventions. These interventions were comprised of Readiness Test with Review Sessions and Readiness Test with Developmental Software. This study reports the impact these interventions had on student performance in Intermediate I.

This study extends prior literature related to improving performance in Intermediate Accounting by using multivariate models and contrasts to determine the impact of our interventions on student performance in Intermediate I. The authors control for gender, grades earned in each of the two accounting principles courses, whether student took Principles II at school of study or transferred in the course, and overall undergraduate grade point average upon entering Intermediate I.

Results suggest that our interventions did improve student performance in Intermediate I. Average Intermediate I student grades were significantly higher for each of the intervention periods over the Base Period that employed no intervention. The multivariate models' results suggest students with higher grades in Principles II and higher undergraduate grade point averages perform significantly better in Intermediate I, which is consistent with prior research.

This study is based on a natural experiment using naturally occurring data, as we apply interventions to all of our students. A controlled experiment would likely produce cleaner and possibly more valid results for the samples, but would suffer somewhat from generalizability; however, a controlled experiment was not possible in this study. We attempt to overcome limitations of a natural experiment by creating advanced multivariate models with control variables for all relevant characteristics that could impact performance in Intermediate I.

AUTHOR INFORMATION

Terry J. Ward, Ph.D., CPA, Professor of Accounting, and Director of MAcc Program, at Middle Tennessee State University, received his Ph.D. degree in Accounting from The University of Tennessee in 1991 and is certified as a Public Accountant in North Carolina. Dr. Ward joined the MTSU faculty in the fall of 1994. His primary areas of research are financial distress, usefulness of cash flow information, and ordinal logistic regression modeling. He has published forty-three articles, forty-five abstract proceedings, and given numerous presentations at regional and national meetings. He received the first MTSU College of Business Outstanding Accounting Faculty Award in 2000 and was awarded the Deloitte Professorship of Accounting from 2006 to 2013. Dr. Ward also co-founded with Jon Woodroof SEDONA Data Systems in 1995, a web based faculty data management system, and remained co-owner until 2010. Telephone: (615) 898-2341. Email: tward@mtsu.edu

Rebecca A. Foote, M.S., CPA, joined the faculty of the Department of Accounting at Middle Tennessee State University in the spring of 2006, and serves as the Accounting Instructional Coordinator. She received her Bachelor of Science in Home Economics Education from The Ohio State University in 1982 and her Master of Science in Accounting/Information Systems from Middle Tennessee State University in 2005. Ms. Foote is engaged in research relative to the effect of intervention mechanisms on student success in upper division accounting courses. She has received several teaching awards at Middle Tennessee State University, having been named the Most Outstanding Professor in the College of Business four times by the students of Jones College of Business. Ms. Foote also received the MTSU Outstanding Teacher Award in 2011. Telephone: 615-494-8923. Email: rfoote@mtsu.edu

Mary E. Phillips, Ph.D., CPA, received the PhD in Business Administration with a Concentration in Accounting from the University of Kentucky in 2005, and she has taught accounting at the Jones College of Business at Middle Tennessee State University since 2004. Her area of research is financial accounting, and she teaches courses in both

the undergraduate and Masters of Accounting programs, including Principles of Accounting, Intermediate Accounting, Advanced Accounting and Financial Accounting Standards. She holds a CPA certificate in Tennessee and also holds the CGMA certificate. She is a member of the American Accounting Association, the American Institute of CPAs and the Tennessee Society of CPAs, where she serves on the board at the state and local levels. She was awarded the Bridgestone Distinguished Assistant Professor Award at the Jones College of Business for 2007-2008, and she has experience in both public and corporate accounting. As a former Fulbright Scholar, she serves on the Fulbright Committee at Middle Tennessee State University, where the school has recently been recognized for success in the number of students who have been Fulbright recipients.

DATA AVAILABILITY: Contact the first author for information about the data.

ENDNOTES

¹The authors also ran the models keeping the repeat students in the sample and adding a control variable, INT1ATT = number of times student took Intermediate I during the period of observation, to control for the repeat observations. The results were similar to those reported in the study but not as strong. So, we only report the results for the models with the repeat observations deleted. For students who repeated Intermediate I, the period of selection, is based on the type of intervention the student had (or none at all for the Base Period students) during the period in which the student received an official grade.

²The authors also generated categorical models using Ordinal Logistic Regression (OLR) to further validate the results reported from the ANCOVA models. For the OLR models, the dependent variable (INTGRADE) was treated as an ordinal categorical response variable. Although OLR is seldom used in accounting education research, one can argue that it is the most valid statistical method to use to generate regression results in cases where the response is categorical (Agresti 1990). OLR bases its results on frequency counts instead of analyses of means and is not limited by many of the assumptions in linear regression methods (such as normality, equal variances, etc.). Thus, OLR produces valid results even in situations where some cells may contain limited observations. Agresti (1990) discusses the various ordinal logistic regression models, and Hosmer and Lemeshow (1989), as well as Agresti, provide thorough discussions of all loglinear (including regression) designs. The Proc Logistic command, with Class statement, in SAS (2009) was used to generate the logistic models. All results using OLR were the same as those for the ANCOVA models. Thus, for simplicity sake, the authors only report the results for the ANCOVA models.

REFERENCES

1. Agresti, A. 1990. *Analysis of Ordinal Categorical Data*. (New York, NY: John Wiley and Sons).
2. Buehlmann, D. M. 1975. ISU's use of the AICPA's College Testing Program. *Journal of Accountancy* (May) 93-96.
3. Burdick, R. and B. N. Schwartz. 1982. Predicting grade performance for Intermediate accounting. *Delta Pi Epsilon Journal*. 117-127.
4. Burnett, R. D., L. Xu, and S. Kennedy. 2010. Student self efficacy in intermediate accounting: A tool to improve performance and address accounting change. *The Accounting Educators' Journal* (Vol. 20) 109-134.
5. Carlan, P. and F. Byxbe. 2000. Community colleges under the microscope: An analysis of performance predictors for native and transfer students. *Community College Review* (Vol. 28, No. 2) 27-43.
6. Carrington, L. G. 2012. Ready for intermediate accounting? Factors affecting student preparation. *The 2012 Orlando International Academic Conference*. 707-715.
7. Colley, J. R., A. G. Volkan, M. Drucker, and M. A. Segal. 1996. Evaluating the quality of transfer versus nontransfer accounting principles grades. *Journal of Education for Business* (Vol. 71, No. 6) 359-362.
8. Danko-McGhee, K. and J. C. Duke. 1992. Predicting student performance in accounting classes. *Journal of Education for Business* (Vol. 67) 270-274.
9. Davidson, R. A. 2002. Relationship of study approach and exam performance. *Journal of Accounting Education* (Vol. 20) 29-44.

10. Delaney, P. R., D. E. Keys, C. L. Norton, and J. R. Simon. 1979. An admission test for Intermediate Accounting. *The Accounting Review* (January) 155-162.
11. Doran, B., M. Bouillon, and C. Smith. 1991. Determinants of student performance in Principles of Accounting I and II. *Issues in Accounting Education* (Vol. 6) 74-84.
12. Eckel, N. and W. A. Johnson. 1983. A model for screening and classifying potential accounting majors. *Journal of Accounting Education* (Fall) 57-65.
13. Eikner, A. and L. Montondon. 2001. Evidence on factors associated with success in intermediate accounting I. *Accounting Educators' Journal* (Vol. 13) 1-17.
14. Fogarty, T. and P. Goldwater. 2010. Beyond just desserts: The gendered nature of the connection between effort and achievement for accounting students. *Journal of Accounting Education* (Vol. 28, No. 1) 1-12.
15. Frakes, A. H. 1977. Introductory accounting objectives and Intermediate Accounting performance. *The Accounting Review* (January) 200-210
16. Gist, W., H. Goedde, and B. Ward. 1996. The influence of mathematical skills and other factors on minority student performance in Principles of Accounting. *Issues in Accounting Education* (Vol.11) 49-59.
17. Hicks, D. W. and F. Richardson. 1984. Predicting early success in Intermediate Accounting: The influence of entry exam and GPA. *Issues in Accounting Education*. 61-76.
18. Hosmer, D. W., and S. Lemeshow. 1989. *Applied Logistic Regression*. (Hew York, NY: John Wiley and Sons).
19. Huang, J., J. O'Shaughnessy, and R. Wagner. 2005. Prerequisite change and its effect on Intermediate Accounting performance. *Journal of Education for Business* (Vol. 80) 283-288.
20. Kilpatrick, B. G., K. S. Savage, and N. L. Wilburn. 2011. Supplemental Instruction in the first intermediate accounting course: Investigation of an intervention strategy to improve student performance. Working Paper, Northern Arizona University, December.
21. Laband, D., D. Rosenberg, and K. Smith. 1997. An examination of the performance of transfer versus "native" students in upper-level accounting courses. *Journal of Accounting Education* (Fall) 515-529.
22. Lipe, M. 1989. Further evidence on the performance of female versus male accounting students. *Issues in Accounting Education* (Vol. 4) 144-152.
23. Lynn, S. A. and I. Robinson-Backmon. 2006. Academic success of non-traditional students: Factors affecting performance in an upper-division undergraduate accounting course. *Journal of College Teaching and Learning* (Vol. 3, No. 12) 85-96.
24. McCormick, R. R. and J. B. Montgomery. 1974. The use of AICPA tests in evaluating transfer students at a university. *Journal of Accountancy* (August) 82-88.
25. Mutchler, J., J. Turner, and D. Williams. 1989. The performance of female versus male accounting students. *Issues in Accounting Education* (Vol. 4) 1103-1111.
26. Phillips, M. E., R. A. Foote, T. J. Ward, and P. B. Thomas. 2013. Developing an assessment and development plan for students entering Intermediate Accounting I: The process and student reaction to the plan. *Academy of Educational Leadership Journal* (Vol. 17, No. 4) 27-42.
27. Sanders, D. E. and V. F. Willis. 2009. Setting the P.A.C.E. for student success in Intermediate Accounting. *Issues in Accounting Education* (Vol. 24) 319-337.
28. SAS Institute Inc. 2009. *PC SAS/STAT*. (SAS Institute Inc.).
29. Shoulders, C. D. and S. A. Hicks. 2008. ADEPT learning cycles enhance Intermediate Accounting student learning success. *Issues in Accounting Education* (Vol. 23) 161-182.
30. Turner, J. L., S. A. Holmes, and C. E. Wiggins. 1997. Factors associated with grades in intermediate accounting. *Journal of Accounting Education* (Vol. 15, No. 2) 269-288.
31. Tyson, T. 1989. Grade performance in introductory accounting courses: Why female students outperform males. *Issues in Accounting Education* (Vol. 4) 153-160.
32. Waples, E. and M. Darayseh. 2005. Determinants of students' performance in intermediate accounting. *Journal of College Teaching and Learning* (Vol. 2, No. 12) 87-92.