Volume 10, Number 3 **ISSN 1095-6328**

ACADEMY OF EDUCATIONAL LEADERSHIP JOURNAL

An official Journal of the

Allied Academies, Inc.

Michael Shurden and Royce Caines
Editors
Lander University

The Allied Academies, Inc., is a non-profit association of scholars, whose purpose is to support and encourage research and the sharing and exchange of ideas and insights throughout the world. Academy information is published on the Allied Academies web page at www.alliedacademies.org.

Whitney Press, Inc.

Printed by Whitney Press, Inc.
PO Box 1064, Cullowhee, NC 28723
www.whitneypress.com

Authors provide the Academy with a publication permission agreement. Allied Academies is not responsible for the content of the individual manuscripts. Any omissions or errors are the sole responsibility of the individual authors. The Editorial Board is responsible for the selection of manuscripts for publication from among those submitted for consideration. The Publishers accept final manuscripts in digital form and make adjustments solely for the purposes of pagination and organization.

The *Academy of Educational Leadership Journal* is published by the Allied Academies, Inc., PO Box 2689, 145 Travis Road, Cullowhee, NC 28723, (828) 293-9151, FAX (828) 293-9407. Those interested in subscribing to the *Journal,* advertising in the *Journal,* submitting manuscripts to the *Journal*, or otherwise communicating with the *Journal*, should contact the Executive Director at info@alliedacademies.org.

ACADEMY OF EDUCATIONAL LEADERSHIP JOURNAL
EDITORIAL BOARD MEMBERS

ACADEMY OF EDUCATIONAL LEADERSHIP JOURNAL

CONTENTS

LETTER FROM THE EDITORS

Welcome to the *Academy of Educational Leadership Journal.* The *AELJ* is published by the Allied Academies, Inc., a non profit association of scholars whose purpose is to encourage and support the advancement and exchange of knowledge, understanding and teaching throughout the world. The *AELJ* is a principal vehicle for achieving the objectives of the organization. The editorial mission of this journal is to publish empirical, theoretical and scholarly manuscripts which advance the discipline, and applied, educational and pedagogic papers of practical value to practitioners and educators. We look forward to a long and successful career in publishing articles which will be of value to many scholars around the world.

The articles contained in this volume have been double blind refereed. The acceptance rate for manuscripts in this issue, 25%, conforms to our editorial policies.

We intend to foster a supportive, mentoring effort on the part of the referees which will result in encouraging and supporting writers. We welcome different viewpoints because in differences we find learning; in differences we develop understanding; in differences we gain knowledge and in differences we develop the discipline into a more comprehensive, less esoteric, and dynamic metier.

Information about the organization, its journals, and conferences are published on our web site. In addition, we keep the web site updated with the latest activities of the organization. Please visit our site and know that we welcome hearing from you at any time.

Royce Caines and Michael Shurden
Editors
Lander University

ACCOUNTING MAJORS' FINANCIAL REPORTING KNOWLEDGE AND THEIR ABILITY TO IDENTIFY AND CORRECT FINANCIAL STATEMENT ERRORS AND OMISSIONS

Marianne L. James, California State University, Los Angeles

ABSTRACT

Financial statement frauds that have impaired the integrity of financial reporting and led to new regulation (i.e., the Sarbanes-Oxley Act of 2002) have reemphasized the importance of financial reporting knowledge. Accounting majors must not only know the multitude of accounting rules, but also the overall reporting requirements. Yet, students frequently lack some knowledge in this area.

The purpose of this study was to investigate accounting majors' financial reporting knowledge, which is necessary to succeed in the challenging accounting profession, and to develop recommendations that will help accounting educators address any weaknesses. A multi-course case project consisting of financial statements and notes containing intentional errors and omissions was utilized in Intermediate Accounting courses.

The study found that students' identification and correction of errors and omissions varied depending on the specific reporting issue. Overall, a higher percentage of students correctly identified errors and omissions on the face of the financial statements than in the notes. Furthermore, only a relatively small percentage of the students noticed that several comparative financial statement years had been omitted. The results from this study suggest that additional instruction is needed, particularly with respect to overall reporting requirements and the relevance of financial statement note disclosures.

INTRODUCTION

The accounting profession has experienced significant changes during the past few years. Highly publicized accounting frauds involving large companies such as WorldCom and Enron and some surprising audit failures have enhanced the scrutiny of the profession and have led to additional regulation (i.e., the creation of the Sarbanes-Oxley Act of 2002). These recent financial reporting scandals have enhanced cognizance of a long recognized need for high quality and truthful financial reporting.

Accounting professionals must be quite knowledgeable about financial reporting. For example, accounting professionals must be able to prepare financial statements and notes that are free of material misstatements and omissions, and must be able to detect financial statement errors and omissions. Accounting majors should acquire fundamental financial reporting knowledge while completing their accounting curriculum. However, frequently accounting majors experience difficulties preparing financial statements and notes that are relevant, reliable, and in compliance with Generally Acceptable Accounting Principles (GAAP). Accounting educators must utilize the limited time available in financial accounting courses (particularly Intermediate Accounting) to help their students learn and understand the fundamental principles and concepts underlying financial accounting and reporting, as well as the multitude of specific accounting rules that comprise GAAP. An understanding of the strengths and weaknesses of students' financial reporting knowledge is needed for the effective and efficient utilization of this limited class time.

Thus, the purpose of this study was (1) to identify accounting majors' specific strengths and weaknesses regarding financial reporting and (2) to develop recommendations that will help accounting educators address these weaknesses. A multi-course financial reporting project containing errors and omissions was utilized for analysis of students' financial reporting knowledge. Overall, a higher percentage of students were able to identify errors and omissions on the face of the financial statements than in the financial statement notes. Additional discussions regarding the relevance of full disclosures in financial statement notes from the user perspective, and additional time devoted to required note disclosures may be useful in addressing this weakness. Furthermore, additional discussions of overall reporting requirements and reinforcement of basic concepts may be needed.

BACKGROUND LITERATURE

The accounting profession has experienced significant change during the past few years. Large, highly publicized accounting frauds have led to enhanced scrutiny of the profession and to additional regulation. The Sarbanes-Oxley Act of 2002 (SOX), which was signed into law on July 30, 2002 was enacted to address some of the accounting and corporate problems that came to light as a result of these accounting fraud cases.

The stated purpose of the SOX is "To protect investors by improving the accuracy and reliability of corporate disclosures made pursuant to the securities laws, and for other purposes." (U.S. Congress, H.R. 3763, 2002). The provisions of the SOX are organized into eleven titles, many of which directly affect financial accounting professionals. For example, consistent with Section 302: "Corporate Responsibility For Financial Reports," the Chief Financial and the Chief Executive officers of a Securities and Exchange Commission (SEC) reporting firm must certify the "...appropriateness of the financial statements and disclosures contained in the periodic report, and that those financial statements and disclosures fairly present, in all material respects, the operations

and financial condition of the issuer." (U.S. Congress, HR 3763, 2002, 302). Furthermore, consistent with Section 401(a): "Disclosures In Periodic Reports; Disclosures Required," all required financial reports prepared consistent with Generally Accepted Accounting Principles (GAAP) must "reflect all material correcting adjustments . . . that have been identified by a registered accounting firm. . . ." (U.S. Congress, HR 3763, 2002, 401a). In addition, quarterly and annual financial reports ". . . shall disclose all material off-balance sheet transactions" and "other relationships" with "unconsolidated entities" that may have a material current or future effect on the financial condition of the issuer." (U.S. Congress, HR3763, 2002, 401a).

High quality financial statements and truthful and thus useful financial reporting can be achieved only if financial statements are both relevant and reliable. While SOX specifically emphasizes the objective of improving financial reporting reliability, it is not a new objective, but one that provides the foundation for financial accounting and reporting and expressly has been stipulated a quarter of a century ago in the Financial Accounting Standards Board's (FASB) conceptual framework. Specifically, Statement of Financial Accounting Concepts No. 2 "Qualitative Characteristics of Accounting Information," (SFAC 2) identified "reliability" and "relevance" as essential ingredients of useful financial statements (FASB, 1980). Reliability is defined as "The quality of information that assures that information is reasonably free from error and bias and faithfully represents what it purports to represent." (FASB, 1980, Glossary). FASB identified verifiability, representational faithfulness, and neutrality as key ingredients of reliability (FASB, 1980). Relevance is defined as "The capacity of information to make a difference in a decision..." (FASB, 1980, Glossary).

Mr. John M. Foster, former FASB member, recently stated that neutral financial reporting continues to represent one of FASB's most important issues. In addition, he also emphasized the importance of neutrality to U.S. capital markets and stated that the efficient allocation of U.S. capital market resources requires that "creditable, reliable and neutral financial information" is available (Foster, 2003).

The basic concepts and principles underlying financial accounting and reporting and the most important specific accounting rules typically are taught in intermediate accounting courses. A sound conceptual understanding of the significance and applicability of these basic concepts and principles tends to help students understand the many detailed rules dealing with specific accounting issues. For example, the full disclosure principle, which specifies that information that is important enough to influence users' decisions and judgement should be reported (Kieso et al., 2004) provides the conceptual basis for many note disclosures. Accrual of costs of good sold, bad debt expense, and warranty expense represent an application of the matching principle, which specifies that all expenses associated with the earning of revenue should be recognized in the same accounting period as the revenue (Kieso, 2004).

Future accounting professionals must be knowledgeable not only about the broad concepts that lead to high quality financial reporting, but must also be familiar with the detailed rules and

regulations. In addition, accounting standards setting and thus financial accounting and reporting may change in the future and become more principles-based, with fewer specific accounting rules. The FASB has issued a proposal "Principles-Based Approach to U.S. Standard Setting" (FASB, 2002) that if adopted will lead to more principle-based accounting standards, which would reduce the existing overload of specific accounting standards and likely require more professional judgment.

Robert Herz, current chairman of FASB perceived the following potential advantages of principles-based standards: According to Mr. Herz, they may (1) allow companies and auditors to exercise professional judgement, which may enhance the professionalism, (2) lead to easier to understand accounting standards, (3) reduce opportunities to utilize "form over substance," (4) reduce the double jeopardy risk, (5) and support the convergence with the International Accounting Standards Board, which currently utilizes a principles-based approach (Herz, 2003). If this proposal is adopted, an understanding of basic concept will become more important than ever for accounting professionals. Consistent with this project, in October 2004, the FASB added a joint IASB-FASB project to its agenda to develop a common conceptual framework to its agenda (FASB, 2005).

Because accounting rules are very complex, students tend to focus on knowing how to calculate and account for certain numbers, and less on overall financial reporting requirements and note disclosures. However, proper and sufficient disclosures are just as important and notes typically convey much needed information to the financial statement user. In addition, after initially learning the basic concepts and principles, students tend to "forget" their significance to specific accounting issues. In his luncheon address at the annual meeting of the American Accounting Association, Mr. Herz stated that a goal of the FASB-IASB conceptual framework project was that "Curriculum development & teaching should focus on the conceptual framework." (Herz, 2005, 17). He also referred to the revised conceptual framework as a "better teaching tool." (Herz, 2005,16).

Furthermore, the Uniform Certified Public Accounting (CPA) Exam has changed and includes problems that require research and judgment in addition to critical/strategic thinking and good communication skills. Specifically, the new computerized CPA exam requires that students are able to research and analyze situations, statements, and standards and to solve interdisciplinary problems (AICPA, 2002). Financial reporting knowledge is necessary to accomplish this objective.

RESEARCH PURPOSE AND HYPOTHESIS DEVELOPMENT

Educators must help students prepare for the challenging accounting profession. A fundamental and comprehensive understanding of basic accounting concepts and principles, overall reporting rules, as well as detailed accounting rules is essential to enable them to participate in preparing or analyzing highly reliable and relevant financial reports. To help educators utilize the scarce time available in intermediate accounting courses, more must be known about the issues and concepts that students are very knowledgeable about, and about those that require additional

instruction. The results of this study will help identify students' strengths and weaknesses in financial reporting.

Repeated application of some basic concepts and accounting issues and reporting rules will tend to reinforce the concepts, lead to deeper understanding of the issues and concepts, and will tend to improve students' ability to identify errors and omissions and the correct treatment and/or supplementation. Thus, hypothesis H1 and H2 state:

H1: Students enrolled in Intermediate Accounting II are more likely to identify overall financial reporting errors and omissions (omitted financial statement years, mathematical errors), than those enrolled in Intermediate Accounting I.

H2: Students enrolled in Intermediate Accounting II are more likely to identify specific financial reporting errors and omissions that are based on the basic financial reporting concepts, than those enrolled in Intermediate Accounting I.

RESEARCH METHODOLOGY

The researcher developed a multi-course financial reporting project that consists of financial statements and financial statement notes of realistic content and length for a consolidated entity. These financial statements and notes contain a total of 28 intentional errors and omissions relating to Intermediate Accounting I and 32 intentional errors and omissions relating to Intermediate Accounting II. Several of the items coincided in both courses. These were compared and tested in hypothesis H1 and H2. Ninety-two students enrolled in three sections of Intermediate Accounting I during the Winter and Summer 2002 quarters and 78 students enrolled in three sections of Intermediate Accounting II during the Fall and Winter 2002 quarters worked on the project for eight weeks while the related subject matter was covered in class. The students were required to assess the overall correctness, completeness, and articulation of (1) the financial statements and notes as a whole and (2) the statements and notes related to specific issues and topics covered in the particular course. Students were asked to address specific accounting issues covered in only their particular course (e.g., a detailed analysis of accounting for inventory would be required of those students enrolled in Intermediate Accounting I and not of those enrolled in Intermediate Accounting II).

A three-step approach was necessary for the students to meet the objectives of the financial reporting project; these were: (1) a review of the financial statements and notes as they are presented in the project; (2) a review of pertinent financial accounting and reporting requirements related to each specific issue; and (3) an assessment of whether the requirements are met. If the requirements were not met, identification of errors and omissions and suggestions regarding necessary changes and supplementation had to be made. These steps helped students develop and enhance their financial reporting knowledge.

During the ninth week of the quarter, students submitted a written report describing the errors and omission that they had identified and suggesting changes and additions that would correct these errors and omissions consistent with GAAP. These student project reports were utilized for this analysis. Students could choose the format for preparing their reports, but were asked to group items consistent with the sequence of course topics and utilize a concise format. Students were not required to identify all errors and omissions in order to earn a high score on the report. Correct responses were summarized and the data was evaluated utilizing descriptive statistics and two-sample t-tests.

RESULTS

Demographics

Sixty-three percent of the students were female and 37% were male. The majority of the students were juniors, with a small percentage of graduate students enrolled in the courses. Because the data was collected utilizing student project reports, no other demographics data (such as age and ethnicity) were collected. Based on past experience, approximately 60-70% of the students enrolled in Intermediate Accounting at this Western university typically work, more than half of the students tend to be of Asian ethnicity, and their age range tends to be widely dispersed.

Study Results

Tables 1 and 2 present descriptions of the financial statement errors and omissions, the specific financial reporting and accounting issues as well as the needed corrections or supplementation, the applicable accounting concepts, and the mean percentage and associated standard deviations of students who correctly identified the errors/omissions and the needed corrections and necessary supplementation. These results are discussed for each course separately. Table 1 presents study results for Intermediate Accounting I and Table 2 for Intermediate Accounting II.

Intermediate Accounting I

Students enrolled in Intermediate Accounting I could identify a total of 28 errors and omissions. The specific topics covered in Intermediate Accounting I and addressed in the project fall under the following broad topics: (The number of errors and omissions for each category are shown in parentheses).

Overall financial statement completeness and integrity (3)
The balance sheet (6)

The income statement and retained earnings statement (13)
The accounts receivable/revenue cycle (4)
Inventory (2)

The mean percent of correct answers to the 28 items varied considerably between financial reporting categories and items.

Table 1: Intermediate Accounting I

Financial Statement Error (E) or Omission (O)	Financial Reporting and Accounting Issue/necessary Correction	Accounting Concepts	Mean Percent Correct n=92 (SD)
Overall Financial Statement Completeness and Integrity			
Cash flow statement for one prior year is missing (O)	Three years should be presented	Completeness; Intra-company comparability	9 (28)
Retained earnings statement for two prior years is missing (O)	Three years should be presented	Completeness; Intra-company comparability	9 (28)
Mathematical errors resulting from incorrect line items or incorrect classifications (E)	Incorrect Financial Statement totals or subtotals.	Financial Statement reliability	56 (50)
The Balance Sheet			
Debt securities due after one year are classified as current (E)	Debt securities due after one year or the operating cycle (whichever is longer) must be classified as non-current	Time period concept; financial statement classification	58 (50)
Available for sale securities expected to be held for more than one year are classified as current (E)	These securities should be classified as noncurrent.	Time period concept financial statement classification	53 (50)
Goodwill is classified incorrectly (E)	Goodwill must be classified as an intangible asset	Financial statement classification	48 (50)
Premium on bonds payable is listed apart from the bonds payable (E)	Premium on bonds payable must be added to the gross bonds payable	Financial statement classification; relevance; representational faithfulness	34 (48)

Table 1: Intermediate Accounting I			
Financial Statement Error (E) or Omission (O)	**Financial Reporting and Accounting Issue/necessary Correction**	**Accounting Concepts**	**Mean Percent Correct n=92 (SD)**
Notes payable was paid off during the current year, yet it was classified as non-current during prior year (E)	The classifications of liabilities; liabilities that are expected to be paid off within one year or the operating cycle (whichever is longer) should be classified as current	Financial statement classification; full disclosure; relevance	21 (41)
The subsection "contributed capital" is missing (O)	Subsections needed on a balance sheet; a contributed capital section will help users readily differentiate between sources of funds.	Financial statement classifications; relevance	49 (50)
The Income Statement			
Net income on income statement does not agree with income shown in retained earnings statement (E)	The amount of net income shown on the income statement must agree with the amount that is added to the retained earnings statement for the same year.	Financial statement articulation	10 (30)
The subsection "other revenues and expense" is missing (O)	Multiple step income statements are organized by the sources of income.	Relevance	49 (50)
Discontinued operations are classified incorrectly and are listed after extraordinary items (E)	Proper classifications of income statement items; discontinued operations must be classified after income from continuing operations and before extraordinary items	Financial statement classifications; relevance	60 (49)
Discontinued items is listed as only one line item (O)	Financial statement detail provided for discontinued items: discontinued operations are broken into two items: (1) operating income from discontinued items and (2) gain or loss on disposal (actual or estimated).	Relevance	24 (43)

Table 1: Intermediate Accounting I			
Financial Statement Error (E) or Omission (O)	**Financial Reporting and Accounting Issue/necessary Correction**	**Accounting Concepts**	**Mean Percent Correct n=92 (SD)**
The tax effect on discontinued items was omitted (O)	Detail provided on discontinued items; tax effect and the net amount must be disclosed.	Relevance	28 (45)
Disclosure on prior year discontinued operations was omitted (O)	Detailed disclosure about a discontinued segment or product line must be provided.	Relevance; full disclosure	5 (21)
Amount shown in notes regarding current year discontinued item does not agree with income statement amount (E).	Financial statement/note articulation. The amount shown in the notes must agree with the amount shown on the income statement.	Articulation between statement and notes; accuracy	3 (17)
Cumulative change in accounting principle is classified incorrectly (E)	Financial statement classification of change in accounting principle. It should be the last item before net income.	Financial statement classification; usefulness	36 (48)
Earnings per share amount is not broken down into all the required components (O)	Earnings per share presentation: EPS amounts must be broken down between EPS from continuing, discontinued, extraordinary, cumulative change in accounting principle and added to EPS from net income.	Full disclosure; relevance	30 (46)
Individual EPS items are not added to EPS from net income (O)	Earnings per share presentation	Full disclosure; relevance	22 (41)
The retained earnings adjustment for the inventory misstatement is incorrect (E)	Accounting treatment for prior year inventory error: No adjustment is made because the error has already counterbalanced	Reliability	7 (25)

Table 1: Intermediate Accounting I			
Financial Statement Error (E) or Omission (O)	**Financial Reporting and Accounting Issue/necessary Correction**	**Accounting Concepts**	**Mean Percent Correct n=92 (SD)**
Incorrect adoption year for SFAS 130 (E)	Compliance with a new mandatory change in accounting principles: According to the effective year, the standard should have been adopted in 1998.	Reliability; full disclosure	1 (10
The note regarding SFAS 123 stated that there was no significant effect; however, the income statement shows and effect of $500,000 (E)	Consistency between financial statements and notes: Given the financial position, $500,000 may be significant to the entity	Materiality; relevance	4 (20)
Accounts Receivable/Revenue Cycle			
Notes disclose that allowance method of estimating uncollectibles is used; the information concerning the percentage estimates has been omitted (O)	Disclosure regarding allowance for uncollectibles: some information concerning the percentage estimates is needed	Relevance; full disclosure	7 (25)
Notes indicate that warranties are accrued when customer returns merchandise (E)	Timing of recognition of warranty related expense: warranty expense must be accrued during year of sale as operating expense	Accrual accounting; timing or expense recognition; relevance; matching	16 (36)
An estimate of warranty cost was not provided (O)	Estimation of future obliga-tions related to warranty on sales: an esti-mate of warranties must be provided	Full disclosure; recognition; materiality; relevance	7 (25)
The net accounts receivable balance was not shown on the face of balance sheets (O)	The net accounts receivable balance must be shown on face of balance sheets	Realizability; relevance	18 (38)
Inventory			
The note indicates that FIFO and LIFO inventory cost flow assumptions are utilized; no disclosure is provided with respect to the type of inventory that is carried at LIFO and what type of inventory is carried at FIFO (O)	Disclosure when several GAAP inventory methods are utilized: disclosure is needed in respect to the type of inventory that is carried at LIFO and FIFO.	Full disclosure; relevance	10 (30)

Table 1: Intermediate Accounting I			
Financial Statement Error (E) or Omission (O)	**Financial Reporting and Accounting Issue/necessary Correction**	**Accounting Concepts**	**Mean Percent Correct n=92 (SD)**
Gain from insurance proceeds from damaged inventory was not classified as an extraordinary item	Classification of gain from insurance proceeds -damaged inventory. A gain or loss from arising from insurance proceeds from damaged inventory is considered unusual and infrequent and must be classified as an extraordinary item and shown net of tax, listed between discontinued items and cumulative change in accounting. principles	Income statement classification; representational faithfulness.	51 (50)

Overall Financial Statement Completeness and Integrity

The students were asked to assess the overall completeness and integrity of the financial statements. The applicable accounting concepts were completeness and intra-company comparability, and reliability.

The set of financial statements, which consisted of two comparative balance sheets, two statements of cash flows, and one statement of retained earnings were incomplete. In addition, several totals or subtotals were incorrect because of incorrect classifications and statement items. Correct identification of these omissions was quite low. Only 9% of students enrolled Intermediate Accounting I identified the omissions of one comparative year for the statement of cash flows and two years of the statement of retained earnings. Fifty-six percent of the students were able to identify the mathematical errors.

Balance Sheet

On the balance sheet, five items were misclassifed within major categories and within time periods (current, non-current). In addition, one major subsection category (contributed capital) had been omitted. The mean percentage of correct identification of errors and omissions varied from 21 to 58 percent between different classification items, with the highest percentage associated with the correct classification of investment securities and the lowest associated with the classification of

liabilities. Approximately half of the students noticed that the subcategory "contributed capital" was needed.

The Income Statement and Retained Earnings Statement

On the income statement, net income did not agree with the net income shown on the statement of retained earnings, violating the concept of financial statement articulation. Furthermore, five errors/omissions were related to the accounting for and disclosure of discontinued operations, violating concepts of full disclosure, relevance, articulation between statements and notes, accuracy and usefulness. In addition, two items were related to the presentation of earnings per share, violating the concepts of full disclosure and relevance. Further, the cumulative change in accounting principle was categorized incorrectly, impairing financial statement usefulness, and an income statement subcategory was omitted, impairing relevance. Furthermore, two items related to the adoption and disclosure of Statements of Financial Accounting Standards Numbers 123 and 130, impairing the concepts of reliability, relevance, materiality, and full disclosure.

Mean correct responses varied considerably. The lowest percent related to the proper adoption year for SFAS 130 (1%), the omitted disclosure of a prior year discontinued operations (5%), note and statement articulation for the discontinued operations (3%), and the articulation problem between the income and retained earnings statements (10%). The highest percentage was associated with the proper financial statement categorization for the discontinued operations (60%), the omission of an income statement subcategory (49%), and the proper categorization of the cumulative change in accounting standards (36%).

Accounts Receivable/Revenue Cycle

Of the four error/omission items two pertained to the disclosures and presentation of accounts receivables and two to accounting for and disclosure of warranty expenses. Seven percent of the students correctly identified the insufficient disclosure of the allowance method, while 18% of the students identified the incorrect presentation of net accounts receivables on the balance sheet, impairing full disclosure, materiality, and relevance. Sixteen percent identified the incorrect description of the proper timing of recognition of warranty expense that violated the matching principle and relevance. In addition, only 7% recognized that some relevant disclosures regarding warranties were omitted, violating full disclosure and relevance.

Inventory

Two items related to an omitted relevant inventory disclosure and the incorrect financial statement classification of a gain on an inventory loss, which violated the concepts of full disclosure,

relevance and representational faithfulness. Fifty-one percent of the students correctly identified the misclassifed gain on the inventory loss and 10% identified the insufficient disclosures regarding the inventory costing methods.

Intermediate Accounting II

Students enrolled in Intermediate Accounting II could identify a total of 32 errors and omissions. The specific topics covered in Intermediate Accounting II and addressed in the project fall under the following broad topics: (The number of errors and omissions for each category are shown in parentheses).

Overall financial statement completeness and integrity (3);
The statement of cash flows (7);
Accounting for investments (9);
Accounting for bonds (3);
Accounting for stockholders' equity (6);
Accounting for property, plant, and equipment (4).

The mean percent of correct answers to the 32 items varied considerably between financial reporting categories and items.

Table 2: Intermediate Accounting II

Financial Statement Error (E) or Omission (O)	Financial Reporting and Accounting Issue/necessary Correction	Accounting Concepts	Mean Percent Correct n=78 (SD)
Overall Financial Statement Completeness and Integrity			
Cash flow statement for one prior year is missing (O)	Three years should be presented	Completeness; Intra-company comparability	51 (50)
Retained earnings statement for two prior years is missing (O)	Three years should be presented	Completeness; Intra-company comparability	30 (46)
Mathematical errors resulting from incorrect line items or incorrect classifications (E)	Incorrect Financial Statement totals or subtotals	Financial Statement reliability	69 (47)
The Statement of Cash Flows			
The increase in wages payable was shown incorrectly as $50,000 (E)	The effect of changes in wages payable on cash from operations. The correct change was $80,000.	Verifiability; financial statement articulation; accuracy	71 (46)

Table 2: Intermediate Accounting II			
Financial Statement Error (E) or Omission (O)	**Financial Reporting and Accounting Issue/necessary Correction**	**Accounting Concepts**	**Mean Percent Correct n=78 (SD)**
The increase in accounts payable was shown incorrectly (E)	The effect of change in accounts payable on cash from operations must agree with the change shown in balance sheet account	Verifiability, financial statement articulation; accuracy.	51 (50)
Cash flows from investing activities was calculated incorrectly (E)	The effect of various investing activities on total cash from investing activities; the correct amount is 2534.	Verifiability; financial statement articulation; accuracy	44 (50)
The proceeds from issuing bonds was incorrectly stated (E)	Consistent with the notes and the increase in net bonds payable, the amount should be 5,220.	Financial statement and note articulation; verifiability; reliability	40 (49)
The taxes paid were not disclosed (O)	Supplemental disclosure required per SFAS 95 for the indirect method: Income tax paid must be disclosed as a supplemental disclosure to the statement of cash flows	Full disclosure; relevance	60 (49)
The trade-in allowance on an old machine was not netted against the cost of the new machine (E)	The treatment of a trade-in allowance in an asset exchange: the trade-in allowance must be netted against the cost of the new machine.	Cost principle	30 (46)
The goodwill amortization expense was not be added in the calculation of cash from operations under the indirect method (E)	The effect of non-cash transactions on cash from operations: the good-will amortization expense must be added in the calculation of cash from operations under the indirect method	Reliability; cash versus accrual	16 (37)
Investments			
The market value of available for sale investments is incorrectly stated as 1,750,000 (E)	Consistent with SFAS 115, "available for sale" securities are recognized at market value at the end of the accounting period.	Fair market value of financial instruments; relevance	53 (50)
Some of the "held to maturity" investments are incorrectly classified as non-current. (E)	Consistent with the information provided in the notes, $100,000 of the held to maturity investments must be classified as non-current because they are due later than one year from the balance sheet date.	Relevance; financial statement classification	51 (50)
The market valuation adjustment was calculated incorrectly (E)	Calculation of the market value adjustment for available for sale securities	Reliability	14 (35)

Table 2: Intermediate Accounting II			
Financial Statement Error (E) or Omission (O)	**Financial Reporting and Accounting Issue/necessary Correction**	**Accounting Concepts**	**Mean Percent Correct n=78 (SD)**
The goodwill amortization does not appear to be reasonable (E)	Calculation of goodwill amortization; given the information provided in the notes regarding the subsidiaries, the goodwill amortization appears incorrect	Reliability	33 (47)
The notes incorrectly refer to goodwill as the difference between cost and book value (E)	The nature of goodwill; goodwill represent the unidentifiable asset and is the difference between acquisition cost and fair market value of identifiable assets.	Cost principle; intangible assets	11 (32)
The notes incorrectly state that "held to maturity" investments are carried at fair market value (E)	Valuation of held to maturity securities. Consistent with SFAS 115, "held to maturity" investments are carried at amortized cost.	Valuation	11 (31)
Goodwill is not classified as an intangible asset (E)	Classification of an unidentifiable asset; goodwill must be classified as an intangible asset	Financial statement classification	57 (50)
The intercompany accounts receivable and accounts payable were not eliminated (E)	The effect of intercompany transaction on the consolidated financial statements	Representational faithfulness.	24 (43)
The notes state incorrectly that goodwill arises from a pooling of interests business combination (E)	Accounting for the excess of cost over book values under a pooling of interests method. Consistent with APBO 16 (now superceded), good-will is not associated with a pooling of interests; assets and liabilities remain at cost; the excess over cost is reflected in retained earnings and additional paid in capital.	Valuation of assets acquired in a pooling of interests business combination.	9 (28)
Bonds			
The bond premium is shown separately from the gross bond payable (E)	Classification of premium on bonds. The bonds premium must be added to par to show the entire bond liability at net present value	Relevance; representational faithfulness; net present value	26 (44)
The amount of the premium shown on the balance sheet does not agree with the amount shown in the notes (E)	The amount shown in the notes must agree with the amount shown on the balance sheet	Financial statement and note articulation; reliability	27 (45)

Table 2: Intermediate Accounting II			
Financial Statement Error (E) or Omission (O)	**Financial Reporting and Accounting Issue/necessary Correction**	**Accounting Concepts**	**Mean Percent Correct n=78 (SD)**
The notes regarding the bond payable are inadequate; the effective (market) interest rate and the issue date(s), and detail on associated costs are not disclosed (O)	Disclosure regarding bonds; extensive disclosure regarding significant bond payables are required; these include the effective (market) interest rate and the issue date(s), and detail on associated costs.	Full disclosure; relevance	34 (48)
Stockholders' Equity			
The notes are incorrect regarding the additional shares issued and are inconsistent with the increase shown on the balance sheet (E)	Consistency between financial statement and note disclosures	Reliability; articulation	59 (50)
The subsection "contributed capital" is missing (O)	Subsections needed on a balance sheet; a contributed capital section will help users readily determine	Financial statement classifications; relevance	37 (49)
The statement of comprehensive income has been omitted, yet the notes indicate that the applicable accounting standard has been adopted (O)	Consistent with SFAS 130, a statement of comprehensive income must be shown.	Full disclosure; relevance	10 (30)
According to the notes, a stock dividend was granted; this dividend is not reflected in retained earnings (E)	Financial statement effect of a stock dividend: this dividend must be shown as a reduction of retained earnings	Reliability; representational faithfulness	47 (50)
The required proforma statement, showing the effect of the market value of stock options on income, has been omitted (O)	Required disclosures for stock options: consistent with SFAS 123, a company that utilizes the intrinsic value method must provide proforma disclosures assuming the fair market value was used	Full disclosure; relevance	7 (26)
No prior period adjustment was needed; the prior years should have been restated (E,O)	Effect of an error on subsequent years; The error counterbalanced; the prior year financial statements should have been restated	Reliability; full disclosure	23 (42)
Property, Plant and Equipment			
The notes state that gains and losses from disposal were treated as operating income (E)	Classification of peripheral activities: gains and losses from disposal of property, plant and equipment must be classified as "other income."	Representational faithfulness; relevance; financial statement classification	3 (17)

Table 2: Intermediate Accounting II			
Financial Statement Error (E) or Omission (O)	**Financial Reporting and Accounting Issue/necessary Correction**	**Accounting Concepts**	**Mean Percent Correct n=78 (SD)**
The estimated useful life of the depreciable assets was not disclosed (O)	Required disclosures for depreciable assets; the useful life (the period over which the assets are depreciated) must be disclosed	Full disclosure; relevance	14 (35)
The type of assets that comprise property, plant and equipment was not disclosed (O)	Disclosures for property, plant and equipment: The classes of assets that comprise property, plant and equipment must be disclosed	Full disclosure, relevance	20 (41)
According to the notes, organizational expense is being amortized (E)	Accounting treatment of organizational costs; consistent with SOP 98-5, organizational costs are expensed as incurred	Reliability; compliance with GAAP	24 (43)

Overall Financial Statement Completeness and Integrity

Intermediate Accounting II students were asked to assess the overall completeness and integrity of the financial statements. The applicable accounting concepts were completeness, intra-company comparability, reliability, and accuracy.

The set of financial statements, which consisted of two comparative balance sheets, two statements of cash flows, and one statement of retained earnings were incomplete. In addition, several totals or subtotals were incorrect because of incorrect classifications and statement items. Correct identification of these omissions and errors varied, with the highest percentages associated with the mathematical errors in some of the totals and subtotals. Specifically, 69% of the students identified the mathematical errors. Fifty-one percent and 30% respectively were able to identify the missing comparative cash flow statement for one of the prior years and the comparative retained earnings statement for two prior years.

The Statement of Cash Flows

In the operating section of the statement of cash flows, two reconciling items (wages payable and accounts payable) did not agree with the change reflected on the balance sheet. This error violated financial statement articulation and verifiability. Seventy-one percent were able to correctly identify the error regarding wages payable and 51% the error regarding accounts payable. In addition, the goodwill amortization was not added back as required to reconcile net income to cash from operations. This omission, which impairs financial statement reliability and accrual accounting was identified by 16% of the students.

In the investing section of the statement of cash flows, the trade-in allowance associated with the sale of property plant and equipment was omitted from the proceeds, violating the cost principle. This error was identified by 30% of the students. In addition, the total of cash from investing activities was incorrectly added, violating verifiability. This error was identified by 44% of the students.

In the cash from financing activities section, the bond proceeds did not agree with the amount of the increase shown on the balance sheet; this violated financial statement articulation, verifiability and reliability. This error was identified by 40% of the students.

Furthermore, one of the most common supplemental disclosures required under the indirect method - the amount of taxes paid - was omitted; this omission, which was identified by 60% of the students violated the full disclosure principle and impaired the concept of relevance.

Accounting for Investments

Nine errors and omissions pertained to accounting for investments. Correct identification ranged from 9 to 57 percent. The highest percentages were associated with the proper balance sheet classification of goodwill (57%), the proper valuation of investment securities available for sale (53%), and the classification of securities held to maturity (51%). The lowest percentages were associated with three errors in the notes regarding (1) the method for calculating goodwill (11%), violating the cost principle, (2) the proper valuation of held to maturity securities (11%), and the proper valuation of business combinations that had been accounted for as a pooling of interests (9%), violating the principle of financial statement valuation.

Accounting for Bonds

Three errors and omissions relating to accounting for bonds impaired relevance, representational faithfulness, the full disclosure principle, and the articulation between the financial statements and the notes. Thirty-four percent of the students identified that the notes did not disclose sufficient information regarding the bond issue, 27% noticed that the premium shown on the balance sheet did not agree with the amount shown in the notes, and 26% recognized that the gross bonds payable and the premium must be reported together.

Accounting for Stockholders' Equity

Six errors and omissions pertained to accounting for and reporting of stockholders' equity. These errors and omissions impaired relevance, representational faithfulness, financial statement and note articulation, full disclosure, and reliability. The mean percentage of students who correctly identified these errors and omissions and the appropriate corrections and supplementations varied

between 7 and 59 percent. The highest percentages correct were associated with the inconsistency between the notes and the balance sheet information regarding the additional common shares issued (59%) and the omitted deduction of the stock dividend value from retained earnings (47%). Thirty-seven percent noticed that a contributed capital subsection should be added to the balance sheet. Twenty-three percent recognized that no prior period adjustment to the retained earnings balance was necessary for an inventory error that already had counterbalanced. A relatively small percentage of the students recognized that the required statement of comprehensive income consistent with SFAS 130 (10%) and a proforma statement regarding stock options consistent with SFAS 123 (7%) had been omitted.

Accounting for Property, Plant, and Equipment

Four errors and omissions pertained to accounting for property, plant, and equipment violating the concepts of full disclosure, relevance, reliability, financial statement classification, and representational faithfulness. The mean correct responses ranged from 3 to 24 percent. Only 3% of the students noticed that gains and losses from the disposal of equipment had been treaded incorrectly as an operating income item, 14% realized that the estimated useful life of property, plant and equipment must be disclosed and 20% that the types of assets that comprise property, plant, and equipment must be disclosed. Twenty-four percent of the students noticed that organizational expense cannot be amortized and must be expensed.

Tests of Hypotheses

Table 3: Tests of Hypothesis

Hypothesis	Error or Omission	P-value
H1	Omitted statement of cash flows	0.00**
H1	Omitted statement of retained earnings	0.00**
H1	Mathematical errors	0.06
H2	Misclassifed investment - held to maturity - non-current	0.14
H2	Goodwill classification	0.08
H2	Bond premium	0.36

Note: ** significant at $p < 0.01$.

Hypothesis H1 tested whether students enrolled in Intermediate Accounting II were more likely to identify overall financial reporting errors and omissions (omitted financial statement years, mathematical errors), than those enrolled in Intermediate Accounting I. Both the Intermediate I and the Intermediate II students were expected to notice that one comparative year of the statement of

cash flows and two comparative years of the statement of retained earnings had been omitted. Two-sample t-tests showed that a significantly higher percentage of the students enrolled in Intermediate II noticed these omissions ($p<0.01$) than did the students enrolled in Intermediate I. Surprisingly, students enrolled in Intermediate II were not more likely than those enrolled in Intermediate I to notice mathematical errors in the statements.

Hypotheses H2 tested whether students enrolled in Intermediate Accounting II were more likely to identify specific financial reporting errors and omissions that violate basic financial reporting concepts, than those enrolled in Intermediate Accounting I. Three items related to both Intermediate Accounting I and II. These were misclassified goodwill, an omitted contributed capital subcategory, and the separate recognition of the bond premium apart from the payable. Contrary to expectation, students enrolled in Intermediate Accounting II did not identify a significantly higher percentage of these errors than those enrolled in Intermediate I. Thus, knowledge and understanding of balance sheet classifications and the relevance of these classifications do not appear to be significantly higher for students enrolled in a more advanced course.

CONCLUSIONS AND RECOMMENDATIONS

Students' ability to recognize and correct errors and omissions varied considerably depending on the particular accounting issues and concepts. Students' knowledge tended to be stronger with respect to mathematical errors, financial statement classifications - both with respect to the distinction between current and non-current items and within financial statement categories, and omitted financial statement subcategories. Students' knowledge and understanding in both courses were weakest with respect to errors and especially omissions in financial statement notes, and disclosures relating to new or pending legislation.

The percentage of students who noticed that comparative financial statement years were omitted was disappointingly low, although a significantly higher percentage students enrolled in Intermediate Accounting II identified these omissions. Educators play a vital role in preparing accounting majors for a challenging and rewarding career.

Additional emphasis in intermediate accounting classes is needed to address these weaknesses. Students in both courses should routinely be exposed to comparative financial statements, and discussions of the importance of financial statement note disclosures should be enhanced. Furthermore, the type and detail of relevant and usefulness accounting information that users need for informed decision making should be discussed in class. This could be facilitated by utilizing short exercises or cases that require that students assume the role of investor or creditor and derive the type of information that they would perceive as relevant. This exercise could then be followed by a discussion of the required disclosures for each major accounting topic, including a discussion of how each disclosure meets the information needs of the user.

Furthermore, whenever possible, a new specific accounting topic should be related back to the fundamental accounting concepts (e.g., relevance, reliability, matching, full disclosure, articulation). This will tend to help students understand that specific accounting rules are not discrete rules, but tend to complement the basic conceptual framework. This approach will become even more useful as accounting standard setting may become more principles driven and the accounting profession continues to address environmental and regulatory changes. Students enrolled in Intermediate Accounting II were better able to identify lack of completeness (i.e., missing years) than were students enrolled in Intermediate Accounting I, but were not more likely to correctly evaluate mathematical accuracy and thus financial statement reliability.

REFERENCES

American Institute of Certified Public Accountants (2002). Uniform CPA Examination. Examination Content Specifications. Retrieved August 31, 2005 from http://www.cpa-exam.org/download/CSOs%20for%20revised%20CPA%Exam.pdf.

Financial Accounting Standards Board. (1980). Statement of Financial Accounting Concepts No. 2 Qualitative Characteristics of Accounting Information: Stamford, Conn.

Financial Accounting Standards Board. (2002, October 21). Proposal "Principles-Based Approach to U.S. Standard Setting. Stamford, Conn.

Financial Accounting Standards Board. (2005). Project Updates. Principles-based Approach to U.S. Standard Setting. Retrieved September 4, 2005 from http://www.fasb.org/project/principles-based_approach_project.shtml.

Foster, J. M. (2003). The FASB and the Capital Markets. Financial Accounting Standards Board. Retrieved July 31, 2003 from http://www.fasb.org/articles&reports/ Fosters_FASBReports.pdf.

Herz, R H. (2003, September) Commentary. A Year of Challenges and Change for the FASB. *Accounting Horizons*, 17 (3), 247-255.

Herz, R.H. (2005). The Conceptual Framework Project. 2005 FARS Luncheon. Retrieved September 3, 2005 from http://www.fasb.org/–05_AAA_FARS_Luncheon.ppt.

Kieso, D.E., J.J. Weygandt, & T.D. Warfield. (2004) *Intermediate Accounting,* (11th ed.) John Wiley & Sons, Inc. p. 42.

U.S. Congress. (2002). One Hundred Seventh Congress of the United States of America at the second Session. Sarbanes-Oxley Act of 2002. H.R. 3763.

THE EFFECTS OF PART-TIME INSTRUCTION ON FINAL GRADES IN THE ENGLISH COMPOSITION COURSE AT A COMPREHENSIVE IIA UNIVERSITY

Allen E. Johnson, Christian Brothers University
Sarah T. Pitts, Christian Brothers University
Rob H. Kamery, Christian Brothers University

ABSTRACT

This paper studies the effect that part-time instruction has on students' final grades in the English Composition course at a comprehensive IIA university. We find that part-time instructors assign grades 0.61 points higher than those assigned by full-time instructors. Using a multiple linear regression, in which the response variable is students' grades, the explanatory variable instructor status--i.e., full-time or part-time--is statistically significant at less than a 0.01 level of confidence (p-value approaches zero). Additionally, the explanatory variable GPA is also significant at an alpha level less than 0.01 with p-values approaching zero. The model yielded an adjusted R^2 value of 0.285, indicating that 28.5 percent of students' grades are accounted for by the explanatory variables included in the model.

INTRODUCTION

In the past forty years, attention in higher education has been directed to the problem of grade inflation. One way of theorizing grade inflation uses the consumer metaphor, wherein students are purchasers of grades provided by teaching faculty. Levine and Cureton (1998) eulogize the university conceived of by John Henry Newman by pointing out that contemporary students, unlike those in Newman's university, by and large carry a more diverse load of responsibilities, making convenience a priority in educational choices and contributing to the popularity of suburban campuses and distance education. Dowling (2003) suggests that the rise in grade inflation indicates instructors' complicity in students' attempts to "beat the system" or get the maximum product with the least expenditure of effort. Grade inflation, then, is seen as one result of increased conflict between the university's attempts to correlate student grades with student performance and the inherent negotiation between grade provider and grade consumer. Kamber and Briggs (2004) delineate similarities and differences in grade inflation and economic inflation, identifying grade inflation as a form of exchange. While ideally an indicator of achievement, in a consumerist environment grades also become a method of rewarding students in exchange for positive student

course evaluations (Edwards, 2000). Among the contributing variables found in a comparative grade inflation study was a perceived or actual increase in teacher pressure (Kolevzon, 1981).

Wyles (1998) found that 75 percent of the new hires in higher education were "contingent workers," a statistic consistent with the national trends in business and government of "downsizing, outsourcing and subcontracting." As such, part-time faculty members are more vulnerable to negative student evaluation of instruction. Other studies of this burgeoning population of part-time instructors suggest that they are less engaged in curriculum, instruction, and scholarship at their institutions, as well as being less autonomous and more demanding of students than their full-time colleagues (Rifkin, 1998; Freeland, 1998). Further, Carney, Isakson and Ellsworth (1978) demonstrate that the increased use of student assessment of instructors in salary, promotion, and tenure processes contribute to grade inflation. The implication here is that if the part-time faculty member's position is tenuous, the case for identifying part-time faculty grading practices as a contributing factor to grade inflation as suggested in Sonner (2000) is more compelling. A satisfied consumer, when asked to comment, gives positive evaluations. If a part-time faculty member wishes to solidify a teaching position, navigating the tension between university intentions and student desires by assigning inflated grades may be an effective strategy.

Johnson (2003) suggests that inflationary practices vary across disciplines, from humanities faculty (most lenient) to social science faculty (most neutral) to economics, science, and natural science faculty (most stringent). Cross and Fray (1993) show that instructors adjust grades, i.e. factor non-achievement related issues into grading. Kolevzon (1981) indicates that grade inflation is positively correlated to the use of subjective evaluation methods. Given the results of these studies and the awareness that English Composition instructors typically have a broader exposure to the student body than their colleagues in other disciplines who do not teach undergraduate general education courses, one might expect the trends cited above from across disciplines to impact the grading practices of part-time faculty teaching English Composition.

LITERATURE REVIEW

Studies reporting the link between grades and course evaluations are varied in their results, in part due to the associational meaning of "high grades." Holmes (1971) finds that students who expected high grades were more likely to indicate they learned more and were more interested, challenged, and stimulated by their instructor. However, students expecting lower grades were not more critical of their instructor's performance. Consistent with these results is another study (Schuerich, Graham & Drolette, 1983) indicating that the expected grade is the least predictive factor in a positive student evaluation of faculty, falling well below "helps to understand" and "sincerely interested" in students' evaluative comments. Pascale (1979) finds no correlation between students' knowledge of course grades and students' evaluation of faculty. Garverick and Carter (1962) and Feldman (1976) find no evidence of a bias in students' evaluation of faculty based

on student grades. Over 50 years ago, Anikeeff (1953) found that "grading leniency" was less highly correlated to positive evaluations among upper-division students than among first- and second-year students. Support for the aforementioned consumerist theorizing of student evaluation of faculty practices can be found in Chacko (1983), who finds that student ratings of faculty encourage grade inflation, and Blass (1974), who finds a positive correlation between faculty ratings and student grades. Bausell and Magoon (1972) and Millea and Grimes (2002) connect students' evaluation of faculty ratings to the discrepancy between expected student grades and student GPA, thus linking the issue of distributive fairness to grading leniency. Griffin (2004) finds grading leniency positively associated with good evaluation scores and Kooker (1968) finds *A* students ranking instructors higher than *C* students. While Greenwald and Gillmore (1997) find a correlation between course grades and course ratings, this is qualified by the indication that giving high grades is not a single sufficient producer of good course ratings. Eiszler (2002), tracking faculty efficiency evaluations and students' grades from 1980 to 1999, finds that both have trended upward. Neath (1996), reviewing the research, counsels faculty to grade leniently if they desire improved student evaluations. Boretz (2004), however, indicates that the research does not provide an "iron clad" link between the two variables. McSpirit and Chapman (2000), interpreting open-ended faculty comments on grade inflation, suggest that these contradictions in the research community on the issue reflect the contradictions among instructors.

Several studies indicate that student grades are related to instructor rank (Ford, Puckett & Tucker, 1987; Jackson, 1986; Sonner & Sharland, 1993; Williamson & Pier, 1985). Research on this topic has consistently illustrated that lower-ranking instructors give higher grades than senior instructors do. Bolge (1995) analyzes student learning as a function of instructor status (full-time vs. part-time). Clark (1990) studies the comparison of the achievement of students taught by full-time vs. part-time instructors. Sonner and Sharland (1993) specifically study grading differences between part-time and full-time instructors. They find part-time instructors assign higher grades on average than full-time instructors. Sonner's (2000) study tests the hypothesis that there is no difference in the average grade assigned by part-time and full-time instructors, finding that adjuncts give higher grades. This suggests that part-time instructors assign students higher grades than full-time instructors do, contributing to grade inflation. MacFarland (1997) finds that part-time instructors award a significantly greater frequency of successful grades (*A*, *B*, *C*, or *Pass*) than do full-time instructors. Landrum (1999) finds that higher grades in courses taught by part-time instructors lead students to develop unrealistic expectations about the grades they should receive and lead to pressure on full-time instructors to follow, or risk student reprisals and poor student evaluation ratings. Grenzke (1998) finds that part-time instructors are more likely to be frequently evaluated than are full-time instructors. Jackson (1986) reports that, in general, students do not rate part-time instructors as highly as full-time instructors. On the other hand, Salmons (1993) finds that students expecting to receive *F* grades from part-time instructors significantly lower their instructor's evaluation scores. Those expecting *A* or *B* grades significantly raise their instructor's

evaluation scores. The findings suggest that part-time instructors receive higher ratings, in part, because they are more lenient in their grade assignments. It would seem reasonable to assume that part-time instructors give higher grades than full-time instructors.

Several studies examine other factors that affect student performance. Schmidt (1983) analyzes the relationship between the amount of time a student devotes to studies and his or her subsequent performance. Additionally, many studies examine the relationship between student attendance and performance. Park and Kerr (1990) and Romer (1993) find a positive relationship between attendance and performance. Durden and Ellis (1995) provide a more detailed study of the attendance/student performance relationship by incorporating absenteeism in their analysis. Esco, Julian and Way (1997) and Melnikova, Williams, Pitts and Kamery (2005) study the correlation between perceived course difficulty and withdrawal rates.

Glasure (2002) finds absenteeism impacts grades negatively. Chan, Shum and Wright (1997) examine the effect of attendance, as well as mandatory attendance, on student performance. Controlling for survival biases, Chan, Shum and Wright (1997) find that while student performance improves with attendance, performance is not improved by mandatory attendance. Chan, Shum and Lai (1996) find evidence that a cooperative learning strategy enhances student performance. Similarly, Beets (2003) finds that cooperative learning results in students earning higher grades.

Other factors related to the student also impact grades. For example, Sen, Joyce, Farrell and Toutant (1997) examine the performance of students by segregating their sample into business majors and non-business majors. They find that non-business majors outperform business majors, and they attribute this superiority to better quantitative preparation and higher overall GPAs.

If faculty grading practices include non-achievement related considerations (Cross & Fray, 1993), and if practices associated with grade inflation vary across disciplines (Johnson, 2003), it is reasonable to examine how non-qualitative variables in composition instruction tend to bias toward grade inflation. Bilimoria (1995) cites the lack of objectivity as a key reason for grade inflation. Many practices in first-year composition instruction, e.g. holistic grading and personal narrative assignments, suggest that objectivity may be lost in the process of composition grading. Brookhart (1998) attributes inflation to the conflation of judge (student performance evaluator) and advocate (champion of non-performance related behaviors) roles in grading, producing a mix of quantifiable standards in evaluation and subjective evaluation. Thus it is not surprising that in the Cross and Fray (1993) survey of 365 faculty on grading practices, 50 percent included class participation in their scoring, 29 percent would lower a borderline grade if a student were negative toward the class, and 43 percent would similarly penalize a student who disrupted the class.

Freedman's (1979) research identifies the elements of composition that most affect instructor grading. Thirty-two student essays were rewritten varying the strength of content, organization, and mechanics to produce 12 versions and 96 different papers. Twelve faculty judges graded the essays. Freedman finds that the less objective components--content and organization--had a more significant effect on positive teacher grading than mechanics and punctuation.

Hake (1986), in a similar study, trained judges to grade two sets of compositions, one set of pure personal narratives and the other of expository essays that used personal narrative as a development strategy. The personal narratives were graded on what Hake identified as less objective criteria. Hake observes that writers of pure narratives were evaluated on the extent to which they fit the grader's bias.

Hayes, Hatch and Silk (2000) discern a halo effect in composition grading. To determine consistency in holistic grading, 796 essays were gathered from 13 first-year composition classes at two institutions. Six judges evaluated the essays, and their grades were compared to the grades given by the instructors of the courses. Instructors tended to give essays assigned later in the course higher grades, while judges' grading indicated little difference in quality between essays produced early in the term and those produced later. The mixing of criteria documented in these studies instantiates what Brookhart (1998) characterizes as confusion of roles.

To the conflation of roles one can add the prevalent notion that first-year English Composition coursework is "less important, less prestigious, and less highly compensated"; the difficulty the specialist in literature experiences in dealing with large numbers of students who are novices to academic culture; and the feeling that as a service course, first-year composition has less to do with faculty perceptions of the discipline of writing than it does with university community needs (Roemer, Schultz & Durst, 1999). Further, the aforementioned conflict between faculty power to grade and student power to evaluate is complicated by the power exercised in grading student writing. As Greenberg (1998) asserts in a review of three works on assessment, not only do graders label students, but English Composition graders are asserting the superiority of their perception of writing competency over that of their students.

While the literature points to several factors contributing to grade inflation, this study analyzes the impact of instructor rank--as it pertains to full-time or part-time employment status--on student grades in the English Composition course. Does enrollment in an English Composition course taught by a part-time instructor improve a student's ability to get a higher grade than enrollment in the same course taught by a full-time instructor? In addition, we examine several student characteristics in order to determine whether those variables interact with instructor status.

METHODOLOGY AND RESULTS

Data were collected from all sections of the English Composition course taught at a private comprehensive IIA university. Seven full-time and five part-time instructors were employed to teach the course. Part-time instructors were classified as adjunct faculty. For the sample of 3,017 students, the following data, which we believe to include explanatory factors for student grades, were obtained:

1) The dependent variable, grade in the English Composition course (A, B, C, D, F)
2) The independent variable, status of the instructor (full-time or part-time)
3) The independent variable, status of the student (day or evening student)
4) The independent variable, student major
5) The independent variable, student gender
6) The independent variable, student age
7) The independent variable, student class standing (freshman, sophomore, junior, senior)
8) The independent variable, student GPA

The dependent variable, *grade*, which is recorded on the students' record as an alpha character, was numerically represented in the model as: A = 4.0, B = 3.0, C = 2.0, D = 1.0, and F = 0.0. Although the dependent variable, *grade*, is ordinal data, and since the interval between the grades can be estimated as being ten point intervals (except for the *F* category), the data is considered to closely approximate interval level data. The use of the values 4, 3, 2, 1, and 0 for the letter grades of *A*, *B*, *C*, *D*, and *F* is similar to using the midpoint of a class to estimate descriptive statistics for a frequency distribution. Students who withdrew from the course were deleted from the sample data. Since student withdrawal data was omitted, the results of the study are subjected to survival bias. The lack of control for such bias is recognized as a limitation of the study.

Table 1 examines and compares the sample variances of the grades given by part-time and full-time instructors. Since the F-test value of 36.684 is greater than the F-critical value of less than 1.16 (least value allowed by reference table), it cannot be assumed that the population variances are equal. Thus a two-sample hypothesis test for the equality of population means would employ the t-test, assuming unequal population variances (see Table 2).

Table 1: F-Test: Two-Sample for Variances

	Full-time	Part-time
Mean	2.41	3.02
Std. Deviation	1.168	1.094
Observations	1102	1915
df	1101	1914
F	36.681	
P(F ≤ f) one-tail	~0	
F-Critical one-tail	<1.16	

Table 2 analyzes the relationship between the status of the instructor, i.e., part-time or full-time, and the grade received in the English Composition course. The hypothesis tested was one of no difference in the average grades awarded by part-time vs. full-time instructors (in the population).

The p-value, which approaches zero, represents the probability that both populations, i.e., part-time instructors and full-time instructors, award grades equally. This contention is rejected at any reasonable level of alpha.

Table 2: T-Test: Two-Sample Assuming Unequal Variances		
	Full-time	Part-time
Mean	2.41	3.02
Std. Deviation	1.168	1.094
Observations	1102	1915
Hypothesized Mean Difference	-0.61	
df	2174.659	
t-Stat	-14.175	
t-Critical one-tail	1.645	
P(T t) two-tail	~0	
t-Critical two-tail	1.96	

Several studies have analyzed relationships between student grades and various student characteristics such as age, gender, class standing, attendance on a full-time or part-time basis, and academic major (Chan, Shum & Wright, 1997; Sen, Joyce, Farrell & Toutant, 1997). We decided to include these variables, along with our variable of main concern, i.e., whether the course was taught by a part-time or full-time instructor, and measure their relationships with a multiple linear regression model. In this way, we can analyze the relationship between student grades and the employment status of the instructor (part-time or full-time) while controlling for the various student demographic characteristics mentioned above.

The multiple regression approach will be utilized here (Kamery, Williams & Kugele, 2004). Using the coding method of A = 4 (or 95), B = 3 (or 85), etc., is similar to estimating the mean or standard deviation of data that has been summarized into a frequency distribution. Table 3 presents the results of a multiple regression analysis.

Student major, class standing, day or evening attendance, and student gender were included as indicator variables. None of these indicator variables were significantly related to the grade received. A graphical analysis of the residuals did not indicate serious violations of the model's assumptions. There are no extreme points (outliers). At each grade level, residual variance does not indicate the presence of homoscedasticity; the residuals approximate a normal distribution. The adjusted coefficient of multiple determination shown in Table 3 is equal to 0.285, indicating that 28.5 percent of the change in the dependent variable, grade, is explained by the set of independent variables (which are student characteristics, except for the instructor status variable). The F-

statistic's high value of 151.377 corroborates the existence of a significant relationship between student grades and the set of independent variables.

Table 3: Regression Results

Multiple R	0.535						
R^2	0.287						
Adjusted R^2	0.285						
Standard Error	1.006						
Observations	3017						
ANOVA							
	SS	df	MS	F	Signifi-cance F		
Regression	1072.496	7	153.214	151.377	~0		
Residual	2670.010	2638	1.01				
Total	3742.506	2645					
	Coefficients	Std. Error	t-Stat	Sig.	Order	Correlations Zero-Partial	Part
(Constant)	0.319	0.117	2.713	0.007			
Day or Evening	-0.061	0.083	-0.730	0.465	0.087	-0.014	-0.012
Age	0.004	0.005	0.782	0.434	0.089	0.015	0.013
Gender	0.184	0.040	4.607	0.000	0.114	0.089	0.076
Major	-0.006	0.002	-3.613	0.000	-0.075	-0.070	-0.059
Student Class	0.019	0.030	0.628	0.530	0.107	0.012	0.010
Cumulative GPA	0.523	0.020	26.729	0.000	0.475	0.462	0.440
Instructor	0.555	0.042	13.155	0.000	0.270	0.248	0.216

Independent variables that would be significant at a 0.01 level of confidence include the following:

1) Instructor status (full-time or part-time) t-Stat value = 13.155
2) Grade point average (GPA) t-Stat value = 26.729

None of the other independent variables showed a significant relationship to the course grade.

During the analysis, several issues of interest were identified for possible future research. There was insufficient information derived from this study to explore those issues here. Those issues include the following:

1) Do part-time and full-time instructors employ similar methods of teaching?
2) Do part-time and full-time instructors use similar methods of testing and grading?
3) Is there coverage by part-time and full-time instructors that is consistent with the prescribed courses of study?
4) Is the performance of students in courses that have a writing component different for those students taught by part-time vs. full-time instructors?

CONCLUSION

The primary objective of this paper was to examine the relationship between students' grades in the English Composition course and the employment status of the instructor, i.e., whether part-time or full-time. A multiple regression model, which allowed for the inclusion of many student characteristics, did report a significant relationship between the two factors. We find that a student's cumulative GPA was the strongest predictor of success in the English Composition course. Next in importance was the employment status of the instructor, part-time or full-time. It is recognized that our sample may include selection bias since part-time instructors may teach predominantly at times and places where non-traditional students are enrolled. Our data was collected at a single university; thus, our results may lack universal application.

REFERENCES

Anikeeff, A.M. (1953). Factors affecting student evaluation of college faculty. *Journal of Applied Psychology*, *17*(3), 458-460.

Bausell, R.B. & J. Magoon (1972). Expected grade in a course, grade point average, and student ratings of the course and the instructor. *Educational and Psychological Measurement*, *32*, 1013-1023.

Beets, S.D. (2003). Difference in student learning techniques: Cooperative learning, quizzes, and homework. *Journal of the Academy of Business Education*, *4*, 1-12.

Bilimoria, D. (1995). Modernism, postmodernism, and contemporary grading practices. *Journal of Management Education*, *19*(4), 440-457.

Blass, T. (1974). Measurement of objectivity-subjectivity: Effects of tolerance for imbalance and grades on evaluations of teachers. *Psychological Reports*, *34*, 1199-1213.

Bolge, R. (1995). Examination of student learning as a function of instructor status (full-time versus part-time). ERIC Document Reproduction Service. ED 383, 241.

Boretz, E. (2004). Grade inflation and the myth of student consumerism. *College Teaching*, *52*(2), 42-47.

Brookhart, S.M. (1998). Why "grade inflation" is not a problem with a "just say no" solution. *National Forum*, *78*(2), 3-6.

Carney, P., R.L. Isakson & R. Ellsworth (1978). An exploration of grade inflation and some related factors in higher education. *College and University: The Journal of the American Association of College Registrars*, *53*(2), 217-230.

Chacko, T.I. (1983). Student ratings of instruction: A function of grading standards. *Education Research Quarterly*, *8*(2), 19-25.

Chan, K., C. Shum & P. Lai (1996). An empirical study of cooperative instructional environment on student achievement in principles of finance. *Journal of Financial Education* (Fall), 21-28.

Chan, K., C. Shum & D. Wright (1997). Class attendance and student performance in principles of finance. *Financial Practice and Education* (Fall/Winter), 58-65.

Clark, B. (1990). Comparison of the achievement of students taught by full-time versus adjunct faculty: Governance and management. ERIC Document Reproduction Service. ED 330, 261.

Cross, L.H. & R.B. Fray (1993). College grading. *College Teaching*, *41*(4), 143-149.

Dowling, W.C. (2003). Meaningless grades and the new dishonesty. *Academic Questions*, *16*(4), 57-63.

Durden, G. & L. Ellis (1995). The effects of attendance on student learning in principles of economics. *The American Economic Review*, *5*, 343-346.

Edwards, C. (2000). Grade inflation: The effects on educational quality and personal well being. *Education*, *120*(3), 546-548.

Eiszler, C.F. (2002). College students' evaluations of teaching and grade inflation. *Research in Higher Education*, *43*(4), 483-502.

Esco, G.M., J.D. Julian & P.K. Way (1997). Student withdrawal behavior in introductory college economics course. *The Journal of Economics*, *23*(1), 83-94.

Feldman, K.A. (1976). Grades and college students' evaluations of their courses and teachers. *Research in Higher Education*, *4*, 69-111.

Ford, H., J. Puckett & L. Tucker (1987). Predictors of grades assigned by graduate teaching assistants in physical education. *Psychological Reports*, *60*, 735-739.

Freedman, S.W. (1979). Why do teachers give the grades they do? *College Composition and Communication*, *30*, 161-164.

Freeland, R. (1998). Adjunct faculty in the community college. ERIC Document Reproduction Service. ED 424899, 1-15.

Garverick, C.M. & H.D. Carter (1962). Instructor ratings and expected grades. *California Journal of Educational Research*, *13*(5), 218-221.

Glasure, Y. (2002). Does absenteeism matter in course performance? *Journal of the Academy of Business Education*, *3*(Fall), 32-42.

Greenberg, K.L. (1998). Grading, evaluating, assessing: Power and politics in college composition. *College Composition and Communication*, *49*, 275-297.

Greenwald, A.G. & G.M. Gillmore (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, *52*(11), 1209-1217.

Grenzke, J. (1998). Part-time faculty: Quality issues. ERIC Document Reproduction Service. ED 417, 686.

Griffin, B. (2004). Grading leniency, grade discrepancy and student ratings of instruction. *Contemporary Educational Psychology*, *29*(4), 410-425.

Hake, R. (1986). How do we judge what they write? In K.L. Greenberg and H. Wiener (Eds.), *Writing Assessment Issues and Strategies: Longman Series in College Composition* (pp. 153-167). White Plains, NY: Longman, Inc.

Hayes, J.R., J.A. Hatch & C.M. Silk (2000). Does holistic assessment predict writing performance? *Written Communications*, *17*(1), 3-26.

Holmes, D.S. (1971). The relationship between expected grades and students' evaluations of their instructors. *Educational and Psychological Measurement*, *31*, 951-957.

Jackson, E. (1986). A comparative study to determine the effectiveness of adjunct faculty in the business division at Fayetteville Technical Institute. ERIC Document Reproduction Service. ED 294, 622.

Johnson, V. (2003). *Grade inflation: A crisis in college education*. New York: Springer-Verlag, 235-245.

Kamber, R. & M. Briggs (2004). Grade inflation: Metaphor and reality. *Journal of Education*, *184*(1), 31-37.

Kamery, R.H., R.H. Williams & L.P. Kugele (2004). The relationship between part-time instructors and final grades in principles of business statistics. *Academy of Educational Leadership Journal*, *8*(1), 1-9.

Kolevzon, M.S. (1981). Grade inflation in higher education: A comparative study. *Research in Higher Education*, *15*(3), 195-212.

Kooker, E.W. (1968). The relationship of known college grades to course ratings on student selected items. *Journal of Psychology*, *69*, 209-215.

Landrum, R. (1999). Student expectations of grade inflation. *Journal of Research and Development in Education*, *32*(2), 124-128.

Levine, A. & J.S. Cureton (1998). College life: An obituary. *Change*, *30*(3), 12-19.

MacFarland, T.W. (1997). A comparison of final grades awarded by full-time faculty and part-time faculty by academic center for winter term 1997. Nova Southeastern University, Research and Planning, Ft. Lauderdale, FL. ERIC: ED429482, 1-38.

McSpirit, S. & A. Chapman (2000). Faculty ironies in grade inflation. *Journal of Instructional Psychology*, *27*(2), 75-78.

Melnikova, V., R.H. Williams, S.T. Pitts & R.H. Kamery (2005). Reducing withdrawal rates across disciplines: The special case of economics and mathematics. *The Journal of Economics and Economic Education Research*. (In press)

Millea, M. & P.W. Grimes (2002). Grade expectations and student evaluation of teaching. *College Student Journal*, *36*(4), 582-591.

Neath, I. (1996). How to improve your teaching evaluations without improving your teaching. *Psychological Reports*, *78*, 1363-1372.

Park, K. & P. Kerr (1990). Determinants of academic performance: A multinomial logit approach. *Journal of Economic Education*, *21*(2), 101-111.

Pascale, P.J. (1979). Knowledge of final grade and effect on student evaluation of instruction. *Educational Research Quarterly*, *4*(2), 52-57.

Rifkin, T. (1998). Differences between the professional attitudes of full- and part-time community college faculty [Paper delivered at the American Association of Community Colleges Convention]. ERIC Clearinghouse for Community Colleges. April 27, 1998, Miami, FL, 1-23.

Roemer, M., L.M. Schultz & R.K. Durst (1999). Reframing the great debate on first-year writing. *College Composition and Communications*, *50*(3), 377-392.

Romer, D. (1993). Do students go to class? Should they? *Journal of Economic Perspectives*, *7*(3), 167-174.

Salmons, S.D. (1993). The relationship between students' grades and their evaluation of instructor performance. *Applied Human Resource Management Research*, *4*(2), 102-114.

Scheurich, V., B. Graham & M. Drolette (1983). Expected grades versus specific evaluations of the teacher as predictors of students' overall evaluation of the teacher. *Research in Higher Education*, *19*(2), 159-171.

Schmidt, R. (1983). Who maximizes what? A study in student time allocation. *American Economic Review*, *73*(1), 23-28.

Sen, S., P. Joyce, K. Farrell & J. Toutant (1997). Performance in principles of finance courses by students with different specializations. *Financial Practice and Education* (Fall/Winter), 66-73.

Sonner, B. (2000). A is for "Adjunct": Examining grade inflation in higher education. *Journal of Education for Business*, *76*(1), 5-10.

Sonner, B. & A. Sharland (1993). Grading differences between graduate teaching assistants and faculty members in the introductory marketing class. *Journal of Marketing Education*, *15*(2), 44-49.

Williamson, L. & J. Pier (1985). Relationship between instructor degree status and student grades: A research report. ERIC Document Reproduction Service. ED 264, 632.

Wyles, B.A. (1998). Adjunct faculty in the community college: Realities and challenges. *New Directions for Higher Education* (Winter), *104*, 89-94.

CURRICULA ASSESSMENT USING THE COURSE DIAGNOSTIC SURVEY: A PROPOSAL

William T. Jackson, Dalton State College
Mary Jo Jackson, Dalton State College
Corbett F. Gaulden, Jr., University of Texas of the Permian Basin

ABSTRACT

The purpose of this paper was to develop a model of an alternative approach to assessing courses and ultimately curricula. Borrowing from Job Characteristics Theory, a modified survey, the Course Diagnostics Survey (CDS) was developed. Using this instrument, a model is suggested that measures the attitudes and resulting outcomes at both the course and overall program level. This model suggests a roadmap as to course or program components that directly impact desired outcomes. Hypotheses are suggested to study the potential of the CDS as an appropriate tool for assessment.

INTRODUCTION

Assessment of student learning has moved to the forefront of business schools over the last decade (Palomba & Banta, 1999, 2001; Banta, Lund, Black & Oblander, 1996). Much of this new emphasis is directly attributable to AACSB International expectations. In fact, many perceive this emphasis to increase with the new standards recently adopted that focus on assurance of learning (Black & Duhon, 2003; Mirchandani, Lynch & Hamilton, 2001; Michlitsch & Sidle, 2002).

Most schools seeking to assess student learning fall back on administering standardized tests, imbedding measurements within courses or conducting post surveys. An almost universal approach has been to survey students with either the periodic student evaluations administered each semester or an instrument that is locally prepared that asks a series of questions to provoke attitudinal responses. Many of these instruments lack a high level of internal validity.

The purpose of this research is to suggest a model using a modification of an instrument that has already proven to be valid and reliable at measuring motivational aspects of a job. The instrument to be recommended, the Course Diagnostic Survey (CDS), is adapted from the Job Diagnostic Survey. This instrument addresses what many, such as Charles Duke (2002), see as being as important as actual content absorption-student perceptions. As will be presented in subsequent sections of this paper, the CDS focuses on how the course design creates unique psychological states (student feeling toward their educational environment) and thus creates affective outcomes (satisfaction or lack there of).

BACKGROUND

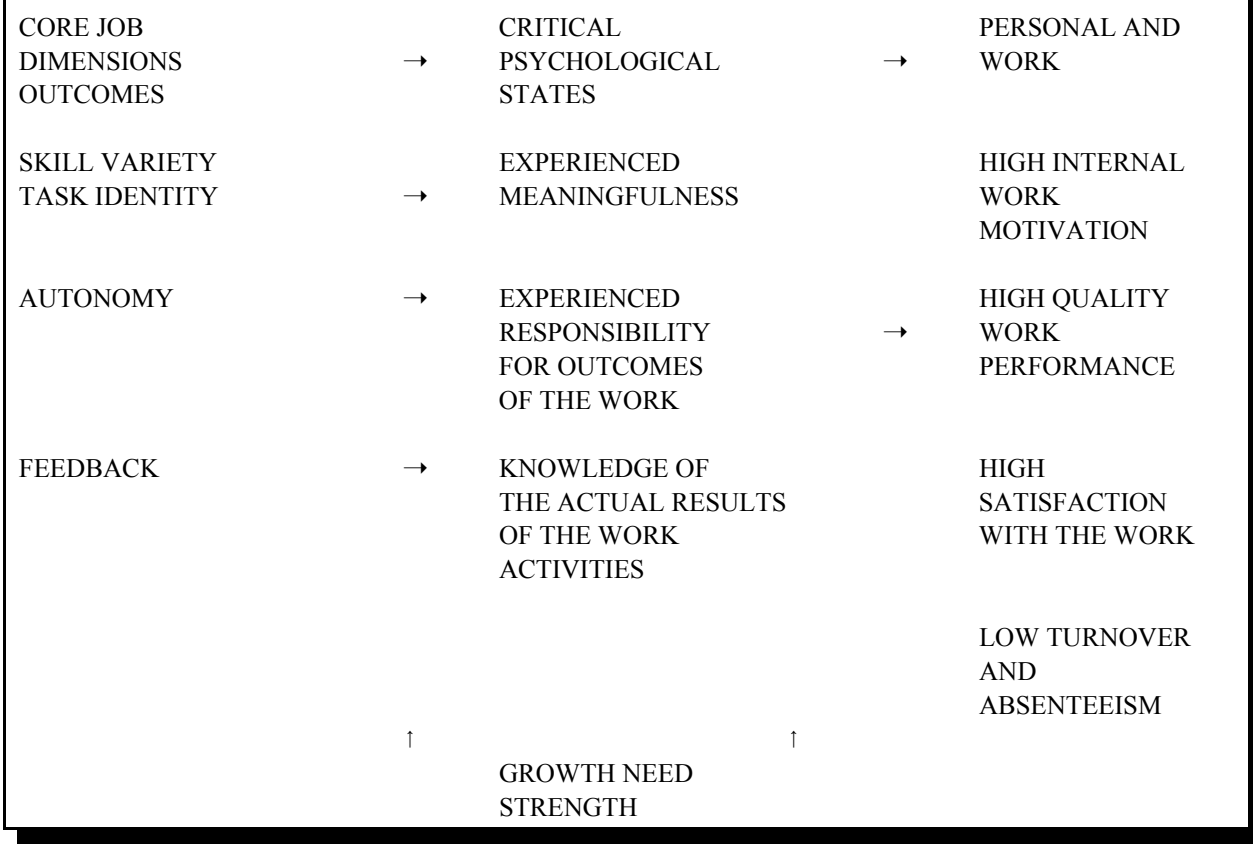
No longer content with just technical competence from our business school graduates, employers are now demanding "…skills in leadership, problem solving, oral and written communication, along with attributes of motivation and assertiveness" (Fontenot, Haarhues & Hoffman, 1991, p. 56). However, the ability of our institutions of higher education to meet these changing demands has been severely questioned. Harvard President Emeritus, Derek Bok (1992) has chided our universities for their failure to even examine the effectiveness of their educational programs.

Fortunately, one stream of research has begun to investigate the effectiveness of selected programs using cognitive scales for this purpose in business curriculums. Using their Skills/Career Usefulness scale, Fontenot, et al. (1991) studied the effectiveness of Small Business Institute (SBI) courses and Business Policy courses in developing desired student skills. Using job analysis and design techniques developed for work environments, Watts and Jackson (1995) investigated the applicability of Hackman and Oldham's (1976) Job Characteristic Theory to course design. Job Characteristic Theory has also been used to assess an institution's student evaluation of instruction (Watts, 1992), and to analyze the effect of course redesign on SBI student outcomes (Watts, Jackson & Box, 1995).

The JDS proposes that positive results will result in the work place (high motivation, high satisfaction with the job and high performance level) when three critical psychological states (experienced meaningfulness of the job, experienced responsibility for the outcomes of the job, and knowledge of the job results) exist. The theory goes on to suggest that the three critical psychological states are created by specific core job characteristics being present. These core job characteristics include: skill variety, task identity, task significance, high levels of autonomy, and effective feedback. However, not all individuals will respond equally, but rather are influenced by their own growth need strength-how important is the job to each person individually. This model is presented below in Figure 1.

As can be seen in Table 1 below, there appears to be an intuitive relationship between what occurs in the job setting to that in the academic classroom. While this relationship may not be exact, it does offer promising possibilities.

FIGURE 1: JOB CHARACTERISTIC THEORY



METHOD

Subjects

The subjects for this study were 586 undergraduate and graduate students in the school of business of a small southwest regional university. Students were represented across all academic disciplines, age distribution, sex, and ethnic background. This number represented nearly 100 percent of all students enrolled in the school. Students were asked to complete the Course Diagnostic Survey. No incentive or penalty was provided for participation in the survey. During this initial phase of the study and for statistical comparison, the results of all participants were combined into one group.

Instrument

As previously mentioned, the instrument used was a modified Job Diagnostic Survey (JDS) resulting is the Course Diagnostic Survey (CDS). The instrument was used to collect perceptions of core course characteristics, critical psychological states, growth need strength, internal academic motivation and course satisfaction. Few modifications were needed to apply the original instrument to the academic environment being examined in this study. Seven point scales was used to maintain consistency with the JDS. This approach has proven to be valid in several other studies involving students in the academic setting (Watts, 1992; Watts, Jackson & Box, 1995; Watts & Jackson, 1995; Fontenot, Haarhues and Hoffman, 1991).

Table 1: Job Characteristics Compared to Course Characteristics

VARIABLE	GENERAL DESCRIPTION	COURSE EQUIVALENCY
SKILL VARIETY	Usage of a wide variety of skills	Usage of a wide variety of skills
TASK IDENTITY	Task closure is evident	Assignments tie together course concepts in a clear manner
TASK SIGNIFICANCE	Outcomes are important	Assignments are important
AUTONOMY	Individuals have impact and are able to make a difference	Students have impact on course outcome
FEEDBACK FROM JOB	Job results are evident	Grades are provided in a timely fashion
FEEDBACK FROM AGENT	Supervisor provides result information	Instructor provides result information independent of grades
MEANINGFULNESS	Work is meaningful	Course is meaningful
RESPONSIBILITY	Responsible for work outcomes	Responsible for course outcomes
KNOWLEDGE OF RESULTS	Final outcomes are known	Final grades are known
GENERAL SATISFACTION	Overall satisfaction with job	Overall satisfaction with course
INTERNAL WORK MOTIVATION	Job is stimulating and challenging	Course is stimulating and challenging
MOTIVATING POTENTIAL STRENGTH (MPS) *	(sk. var. + task id. + task sign.)(autonomy)(job feedback)	(sk. var. + task id. + task sign.)(autonomy)(job feedback)

Course Components

Course components were measured with the seven point scale very inaccurate to very accurate in response to "how much do you agree with the statement". The course component skill variety was measured by the items "the course requires me to use a number of high and complex skills" and a reverse score of "the course is quite simple and repetitive". Task identity was measured by "the course provides me the chance to completely finish the pieces of work I begin" and a reverse score on the item "The course is arranged so that I do not have the chance to do an entire piece of work from beginning to end".

Task significance was indicated through responses on "This course is one where a lot of other people can be affected by how well the work gets" and a reverse response on "The course itself is not very significant or important in the broader scheme of things". Autonomy was shown through the items "The course gives me considerable opportunity for independence and freedom in how I do the work" and "The course denies me any chance to use my personal initiative or judgment in carrying out the work (reverse scored)".

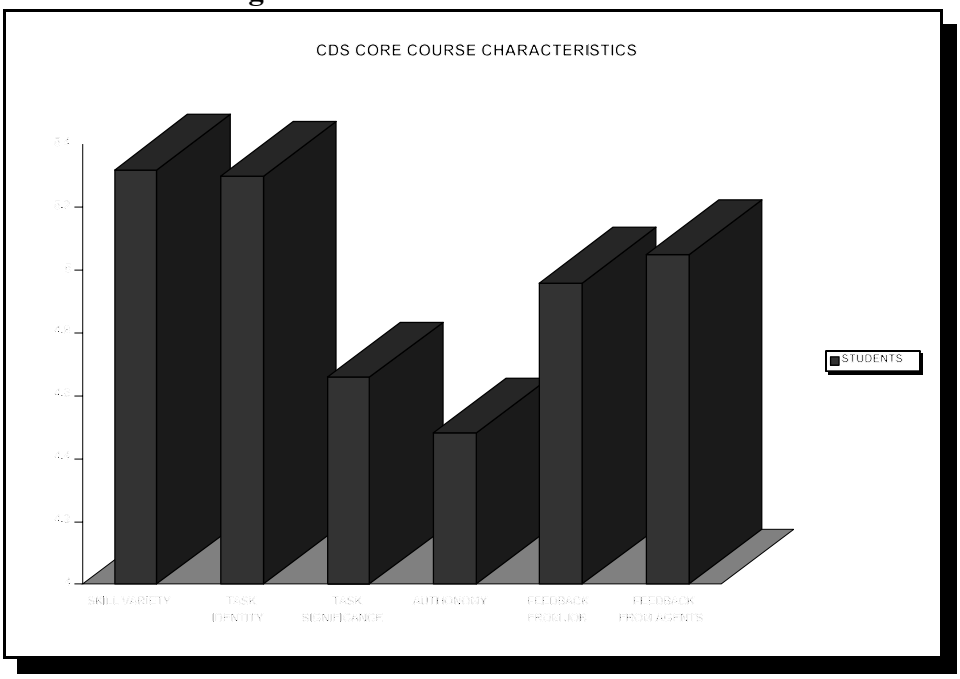
The final two components, feedback from the course and feedback from the instructor, were measured respectively by "Just doing the work required by the course provides many chances for me to figure out how well I am doing", "The course itself provides very few clues about whether or not I am performing well (reversed)" and "The instructor often lets me know how well I am performing", "The instructor and fellow students in this course almost never give me any feedback about how well I am doing in my work (reversed)".

The table and figure below illustrate the means and standard deviations of the sample population for the six course characteristics i.e. skill variety, task identity, task significance, autonomy, feedback from the course, feedback from the instructor.

Table 2: Course Component Means and Standard Deviations

	MEAN	S.D.
Skill variety	5.32	1.25
Task identity	5.30	1.37
Task significance	4.66	1.24
Autonomy	4.48	.86
Feedback from job	4.96	1.37
Feedback from agents	5.05	1.45

Figure 2 Core Course Characteristics



Critical Psychological States

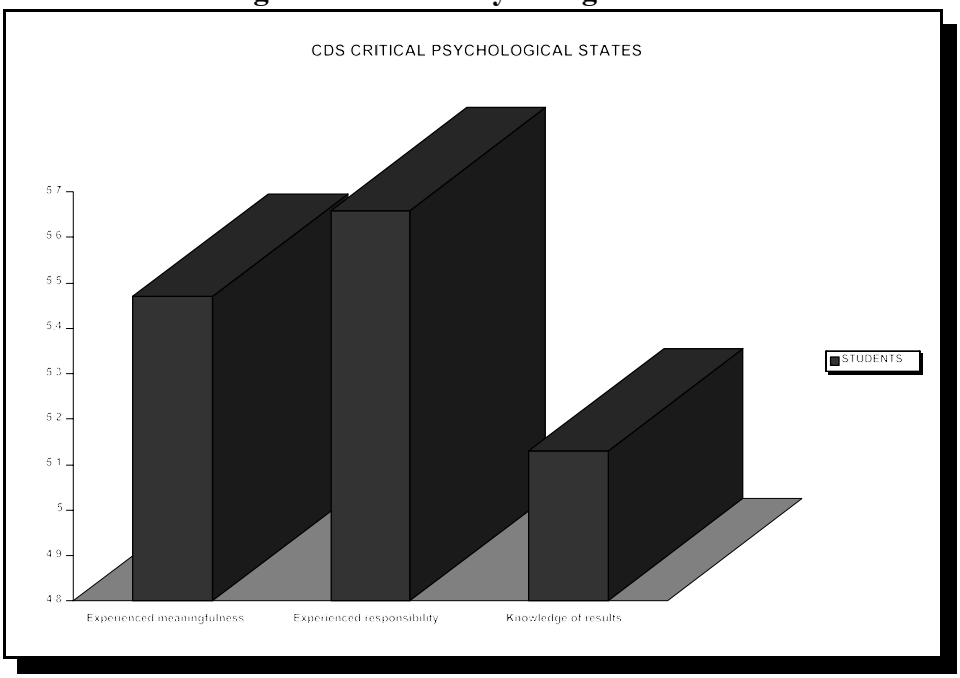
The critical psychological states inspired by the course components were also measured in this study. As stated before, the CDS used a seven point Likert scale. The scale was a measure of how well the student agreed with the statement and scales ranged from very inaccurate to very accurate.

The psychological state meaningfulness was indicated by two items, "The work I do in this course is very meaningful to me" and "Most of the things I have to do in this course seem useless or trivial" which was reverse scored. Responsibility was measured by "I feel a high degree of personal responsibility for the work I do in this course", "I feel I should personally take the credit or blame for the results of my work in this course", "Whether or not course work gets done right is clearly my responsibility" and the reversed item "It's hard, in this course, for me to care very much about whether or not the work gets done right". The last psychological state measured, knowledge of results, was indicated by "I usually know whether or not my work is satisfactory in this course" and "I often have trouble figuring out whether I'm doing well or poorly in this course" which was reverse scored.

The table and figure below indicate the means and standard deviations of the three psychological states, meaningfulness, responsibility, and knowledge of results.

Table 3 Means and Standard Deviations of Psychological States		
Experienced meaningfulness	5.47	1.40
Experienced responsibility	5.66	1.23
Knowledge of results	5.13	1.50

Figure 3 Critical Psychological States



Student Outcomes

The third element component of the CDS was a measure of two student outcomes. These included general satisfaction with the course and student motivation. The seven point Likert scale indicated how well the student agreed with the statement and scales ranged from very inaccurate to very accurate.

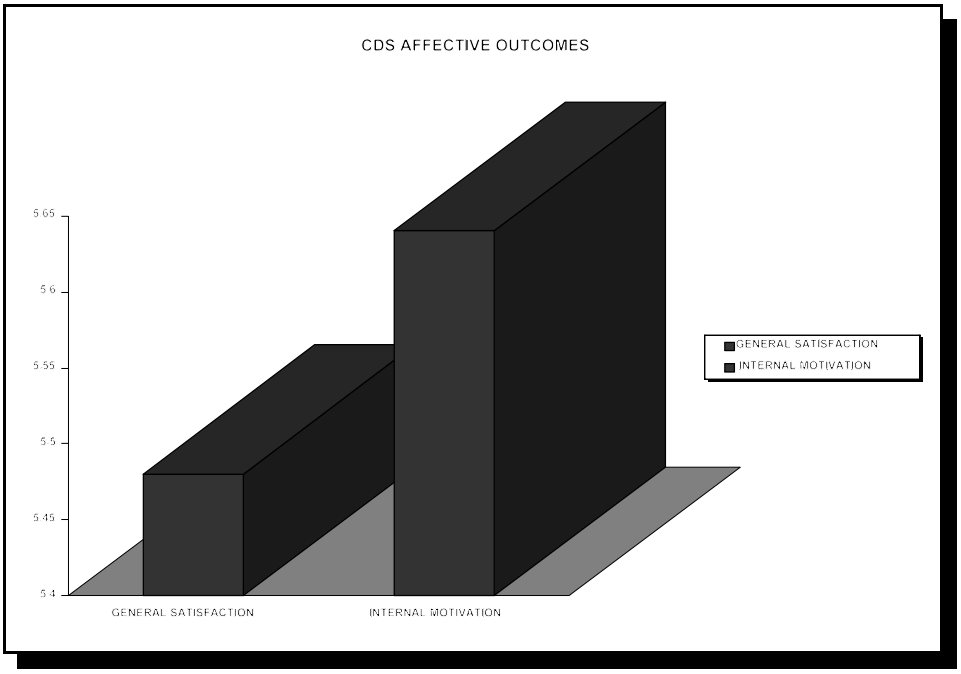
General satisfaction with the course was measured by three items. These included "Generally speaking, I am very satisfied with this course", "I am generally satisfied with the kind of work required in the course" and "I frequently think of dropping this course" that was reverse scored. The outcome, student motivation, was determined by responses on "My opinion of myself goes up when I do the course work well", "I feel a great sense of personal satisfaction when I do the

course work well", "I feel bad and unhappy when I discover that I have performed poorly in this course", and the reverse scored item "My own feelings generally are not affected much one way or the other by how well I do in this course".

The means and standard deviations of the student responses for the general satisfaction and motivation outcomes are shown below.

Table 4 Student Outcomes Means and Standard Deviations		
General satisfaction	5.48	1.54
Internal motivation	5.01	1.01

Figure 4 Student Outcomes



Motivating Potential Score

The Motivating Potential Score indicates the motivating potential of a job or with the CDS, a course. It would be measured by the responses of students in individual courses and calculated by the formula:

$$\text{Motivating potential score (MPS)} = \frac{\text{Skill variety} + \text{Task identity} + \text{Task significance}}{3} \times \text{Autonomy} \times \text{Feedback from the job}$$

PROCEDURE

To capture the influences of course related activities, the instrument was administered late in the semester. On a predetermined date, instructors announced in class that students had been asked to participate in an important study and read the following instructions:

> *This questionnaire was developed as part of a study of course-related activities and how students react to them. The questionnaire helps to determine how courses can be better designed by obtaining information about how students react to different kinds of course-related activities.*
>
> *On the following pages you will find several different kinds of questions about your course. Specific instructions are given at the start of each section; please read them carefully. It should take no more than 15 minutes to complete the entire questionnaire. Please move through it quickly.*
>
> *The questions are designed to obtain your perceptions of course-related activities and your reactions to them. There are no trick questions. Your individual answers will be kept completely confidential. Please answer each item as honestly and frankly as possible.*
>
> *Thank you for you cooperation.*

DISCUSSION

As stated in the introduction, the main purpose of this exploratory study was to propose the use of the Course Diagnostic Survey instrument as a means of assessing students in an academic setting. As the original instrument was intended to assess the impact of redesign of jobs, the modification and resulting CDS could equally be as successful in assessing the impact of specific course components on the psychological states and outcomes in the educational setting.

While it appears that the instrument has potential in this area, additional study is needed to validate the instrument in an educational setting. Specifically, three hypotheses are proposed.

H1: *The CDS course components, i.e. skill variety, skill identity, task significance, autonomy, feedback from course, and feedback from instructor; lead to the indicated psychological states.*

H2: *The critical psychological states in the academic setting as measured by the CDS will be related to general satisfaction and motivation.*

H3: The MPS as measured by the CDS will indicate the motivating potential of a specific course.

Future studies using the study sample, as well as other samples, should attempt to show the relationships indicated in the hypotheses. This would assist in the validation of the CDS and its use in course assessment. It is also suggested that additional studies consider the use of the model in overall program development.

REFERENCES

Ater, E.C. & Coulter, K.L. (1980). Consumer internships: Encouraging consumer/business dialogue. *Journal of Business Communications, 17*(2), 33-39.

Banta, T.W., Lund, J.P., Black, K. & Oblander, F.W. (1996). *Assessment in Practice.* San Francisco: Jossey-Bass Publishers.

Barnett, S.T., Dascher, P.E. & Nicholson, C.Y. (2004). Can school oversight adequately assess department outcomes? A study of marketing curriculum content. *Journal of Education for Business, 79*(3), 157-162.

Black, H.T. & Duhon, D.L. (2003). Evaluation and improving student achievement in business programs: The effective use of standardized assessment tests. *Journal of Education for Business, 79*(2), 90-98.

Bok, D. (1992). Reclaiming the public trust. *Change*, August, 23-29.

Cordery, J.L. & Sevastos, P.P. (1993). Responses to the original and revised job diagnostic survey. *Journal of Applied Psychology, 78*(1), 141-143.

Davis, M.A., Curtis, M.B. & Tschetter, J.D. (2003). Evaluating cognitive training outcomes: Validity and utility of structural knowledge assessment. *Journal of Business & Psychology, 18*(2), 191.

Duke, C.R. (2002). Learning outcomes: Comparing student perceptions of skill level and importance. *Journal of Marketing Education, 24*(3), 203-217.

Fontenot, G., Haarhues, M. & Hoffman, L. (1991). The benefits of the SBI program: Perceptions of former students. *Journal of Small Business Strategy, 2*(1), 56-71.

Hackman, J.R. & Lawler, E.E. (1971). Employee reactions to job characteristics. *Journal of Applied Psychology Monograph, 55*, 259-286.

Hackman, J.R. & Oldham, G.R. (1975). Development of the job diagnostics survey. *Journal of Applied Psychology, 60*, 159-170.

Hackman, J.R. & Oldham, G.R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance, 16,* 250-279.

Hackman, J.R. & Oldham, G.R. (1980). *Work Redesign.* Reading, MA: Addison-Wesley.

Hoffman, L., Fontenot, G. & Viswanathan, R. (1990). An exploratory evaluation of the effectiveness of the SBI program as perceived by quantitative and non-quantitative majors. Proceedings of the 1990 SBIDA National Conference, 80-85.

Kent, T.W. & Davis, T.J. (2002). Using retranslation to develop operationally anchored scales to assess the motivational context of jobs. *International Journal of Management, 19*(1), 10.

Lawrence, E. (1990). Learning portfolio management by experience: University student investment funds. *The Financial Review, 25*, 165+.

Luthans, F. (1992). *Organizational Behavior*. New York: McGraw-Hill, Inc.

Marchese, M.C. (1998). Some factors affecting the relationship between job characteristics and job worth: A job-role interpretation. *Journal of Organizational Analysis, 6*(4), 355-369.

Michlitsch, J.F. & Sidle, M.W. (2002). Assessing student learning outcomes: A comparative study of techniques used in business school disciplines. *Journal of Education for Business, 77*(3), 125-130.

Mirchandani, D., Lynch, R. & Hamilton, D. (2001). Using the ETS major field test in business: Implications for assessment. *Journal of Education for Business, 77*(1), 51-57.

Oldham, G.R., Hackman, J.R. & Stepins, L.P. (1978). Norms for the job diagnostic survey. Technical Report No. 16, School of Organization and Management,Yale University, 1-45.

Palomba, C.A. & Banta, T.W. (2001). *Assessing Student Competence in Accredited Disciplines.* Sterling, VA: Stylus Publishing, Inc.

Palomba, C.A. & Banta, T.W. (1999). *Assessment Essentials*. San Francisco: Jossey-Bass Publishers.

Riggio, R.E., Mayes, B.T. & Schleicher, D.J. (2003). Using assessment center methods for measuring undergraduate business student outcomes. *Journal of Management Inquiry, 12*(1), 68-78.

Rungtusanatham, M. & Anderson, J.C. (1996). A clarification on conceptual and methodological issues related to the job characteristics model. *Journal of Operations Management, 14*(4), 357-367.

Serva, M.A. & Fuller, M.A. (2004). Aligning what we do and what we measure in business schools: Incorporating active learning and effective media use in the assessment of instruction. *Journal of Management Education, 28*(1), 19.

Spicer, D.P. (2004). The impact of approaches to learning and cognition on academic performance in business and management. *Education and Training, 46*(4/5), 194.

Tiegs, R.B., Tetrick, l.E. & Yitzhak, F. (1992). Growth need strength and context satisfactions as moderators of the relations of the job characteristics model. *Journal of Management, 18,* 575-593.

Watts. L.R. & Hudnall, J. (1991). Applying job characteristics theory to course design. In G. James (Ed.), Proceedings of the Mountain Plains Management Conference. Fort Collins, CO: Colorado State University.

Watts, L.R. & Jackson, W.T. (1994). The effect of course redesign on SBI student outcomes: An application of job characteristics model. *Journal of Business & Entrepreneurship, 6*(1), 89-100.

Watts, L.R., Jackson, W.T. & Box, T.M. (1995). Student related outcomes and task design characteristics. *Texas Business Education Association Journal, 5*(1), 127-141.

Watts, L.R. & Jackson, W.T. (1995). The SBI program and student outcomes: A study of business policy classes. *Journal of Small Business Strategy, 6*(1), 93-103.

UNIVERSITY STUDENT ETHICS: THE DIFFERENTIAL EXPLANATORY EFFECT OF LOCUS OF CONTROL

Evelyn C. Hume, Longwood University
Aileen Smith, Stephen F. Austin State University

ABSTRACT

With the ever growing concern over business ethics, an increasing number of business programs require students to take an ethics course. However, researchers have found that ethics instruction alone does not control ethical orientation of students. Individual personal characteristics play a significant role in determining one's ethical decisions and actions. This study examines whether locus of control has discriminating power when the questionable actions are collaborative in nature. Additionally, the research tests whether locus of control has a differential moderating affect when the subjects are considering their own beliefs and actions and when they are considering the actions of others. The findings indicate that the locus of control variable has a significant influence on ethical behavior, even when the actions are collaborative. Moreover, the research shows that locus of control does not significantly influence student's perceptions regarding questionable behaviors of others.

INTRODUCTION

In response to the myriad corporate accounting and financial scandals, colleges and universities have shown a growing commitment to including ethics in the standard business curriculum. Ten years ago, ethics was typically offered as an elective course, if it was included in course offerings at all. Today the number of business programs requiring students to take an ethics course has grown exponentially, and many state societies of CPAs require all accounting students to have ethics instruction to sit for the CPA exam.

Unethical behavior doesn't suddenly arise after one has reached the level of corporate vice president or CEO. Indeed, evidence suggests that cheating may be rampant among students, both at the college and the high school levels. In one recent survey of college students, at least 50% reported cheating on various behaviors (McCabe et al., 2003) and another study indicated a 59% cheating rate (Rael, 2004).á Moreover, an unpublished survey of college freshmen in a small university in the east revealed that some felt cheating was covertly encouraged by their teachers to improve standardized test scores and improve school ratings. There is little doubt that the issue of ethics must be aggressively addressed.

In spite of all the attention given to ethics instruction, many questions about the effectiveness of teaching ethics in the classroom remain unanswered. The extent to which ethics instruction affects ethical orientation is often the object of scientific enquiry, and researchers have found that instruction alone doesn't determine one's ethics. Individual personal characteristics also impact ethical decisions and actions. Locus of control is one such individual characteristic and is the personality variable of interest in this study. Locus of control relates to whether one believes he/she is in control of his/her own destiny. This study examines whether locus of control is a moderating variable when the questionable actions are collaborative activities.

A second issue of interest in this research is whether locus of control has discriminating power when subjects are considering whether questionable actions occur in their environment. Some researchers have asked students how often they believe others engage in unethical behaviors and have used the results as a proxy for the students' own actions (need one or more citations here). However, if there is a difference in the explanatory effect of locus of control between subjects' own actions and their beliefs about the actions of others, researchers must be cautious about interpreting the results that ask about one's own beliefs and actions and the perceived actions of others. In such cases, assumptions drawn from studies using both types of surveys items may be called into question.

LITERATURE REVIEW

Research examining the ethical orientation of university students in the US has been reported in the literature for several decades. The discriminating power of the personality variable Locus of Control (LOC) in ethical responses has been examined by a number of these studies. Rotter's(1966) LOC instrument is designed to assess how much control an individual believes he/she has over the outcomes or consequences in life. LOC is based on whether an individual believes there is a causal relationship between his/her decisions and behaviors and the potential outcomes of those decisions and behaviors.

An ***Internal*** person believes in the causal link between his or her decisions or actions and the expected consequences of those actions. Internals believe that the consequences of their lives are directly related to the decisions they make and the actions they take. In general Internals accept responsibility for what happens to them. ***External*** individuals believe that the expected outcomes or consequences in their lives are not linked to their efforts or decisions. Instead, they believe the outcomes are under the control of luck, fate, or powerful others. Externals generally do not believe in the acceptance of responsibility for what happens to them because their belief structure does not include a cause▲and▲effect relationship between the behavior and the resulting outcome.

Studies of university student ethics have found varying degrees of support for the Internal/External LOC distinction. Some research indicates support for the belief that Internals will supply the more ethical responses to surveys and vignettes (Hegarty& Sims, 1978; Brownell, 1981;

Terpstra et al., 1991; Jones & Kavanaugh, 1996; McCuddy & Peery, 1996; Trevino & Youngblood, 1990; Smith et al., 1999; Beu et al., 2003). Other studies have reported only limited support for the LOC variable effect in ethical research using university students (Hegarty & Sims, 1979; Trevino & Youngblood, 1990; Rogers & Smith, 2001). Research results have also been reported that indicate no ethical response differences on the LOC variable using university students (Brownell, 1982; Geurin & Kohut, 1989; Jones & Kavanaugh, 1996; Bass et al., 1999).

Other research has examined the importance of LOC in whistleblowing activities, job satisfaction, and environmental control contexts. A number of studies have reported general support for the Internal's preference for a participative decision style and greater work satisfaction (Brownell, 1982, 1981;Geurin & Kohut, 1989; Licata et al., 1986; Spector & Michaels, 1986; Storms and Spector, 1987). In an investigation of ethical judgment and whistleblowing, Chiu (2003) reported differences on the LOC variable using non US MBA students. The Internal LOC students indicated more ethical responses then the External Students. These results additionally support the position that Internals expect a cause▲and▲effect relationship between their actions and decisions and the resulting outcomes of those actions and would prefer to participate in the decision process and recognize their stake in the decisions and actions.

RESEARCH HYPOTHESES

This research examines the moderating effect of LOC on students' individual ethical orientation, as well as beliefs about how often the questionable actions occur in their academic environment. Because Internals conceptually link their attitudes and actions with the subsequent results, it is expected that the Internal/External LOC distinction will have an explanatory effect on the students' responses concerning how unethical they believe surveyed questionable behaviors are considered to be.

H_1: Internal LOC students will report more ethically sensitive responses than External LOC students.

The second part of the questionnaire turns the attention from the ethical orientation of the students toward the students' beliefs concerning how often they believe their peers practice the listed questionable behaviors. Several ethical studies surveying students have reported results indicating that they believe their own beliefs and actions are more ethical than their others (Stevens, 1984; Newstrom & Ruch, 1976; Pratt & McLaughlin, 1989; Tyson, 1990).áThe focus of the moderating effect of the LOC variable also changes in the second part of the survey. Because LOC refers to an individual's expected outcomes associated with the actions/decision previously made by that individual, no differences are expected in Part 2 of the survey on the LOC variable.

H_2: There will be no differences in the responses by LOC designation concerning how often the students believe the questionable behaviors are practiced in the students' environment.

METHODOLOGY

The data (n=933) for the analysis were collected from College of Business students from three universities located in the southern US. The data were collected from intact classroom situations, and only six surveys were too incomplete to use in the analysis. The students completed the surveys during class time and were assured both individual and university anonymity. Demographic information on the student respondents is shown in Table 1. The gender breakdown is approximately even, and 63% of the students are upper level or graduate students. In addition, 87% of the student respondents are of traditional age.

Table 1 - Respondent Demographics

Classification:	n*	Percent	Age Classification:	n*	Percent
Freshman	106	13	≤ 20 years old	361	41
Sophomore	204	24	21-25 years old	412	46
Junior	286	34	≥ 26 years old	116	13
Sr/Grad/Spec	241	29	Total	889	100
Total	837	100			
Gender	n*	Percent	Locus of Control	n*	Percent
Male	472	51	≤10 (INT)	504	54
Female	458	49	≥11 (EXT)	429	46
Total	930	100	Total	933	100

* Not all totals equal 933 due to response omissions.

QUESTIONNAIRE

Some of the earlier reported research on ethics using students has used scenarios or questions that are designed to reflect real-world, often business, situations. While these are certainly topics of interest for the study of ethics, university students, particularly undergraduate ones, usually do not have the requisite knowledge or experience to be able to respond appropriately to these types of survey questions. The ethically questionable items surveyed by the current research were selected because they sampled behaviors that are considered to be familiar to the student's in their academic environment. The survey requested that the students respond to their beliefs concerning questionable

behaviors in an academic environment. The behaviors were also generally collaborative and active in nature.

The questionnaire contains four parts. The first two parts were adapted from a questionnaire used by Pratt and McLaughlin (1989). The first part requested from the students information concerning their own personal beliefs toward 14 behaviors of varying degrees of ethicality. The second part of the questionnaire requested that the students respond concerning how often they believed the behaviors occurred in their environment. This part used the same behaviors as Part 1. The third part of the questionnaire used Rotter's (1966) Locus of Control survey to determine the students' LOC scores for the Internal/External classification and analysis. The final section of the questionnaire requested demographic information, which is displayed in Table 1. The gender breakdown is approximately even, and 63% of the students are upper level or graduate students. In addition, 87% of the student respondents are traditional age.

ANALYSIS AND RESULTS

The first two parts of the survey asked the students to indicate their beliefs toward 14 questionable behaviors. Part 1 requested that the students indicate how unethical they believed each of the 14 behaviors to be. A Likert scale was used for the responses where 1 = very unethical and 5 = not at all unethical. The focus of this section of the survey is the ethical orientation of the students toward the academic▲environment behaviors. It should be noted at this point that the behaviors surveyed were of an assumed collaborative or collective nature. The items expressly stated or implied that the individual was acting with one or more other individuals. Part 2 of the survey requested that the students indicate how often they believed most college students practice each of the same behaviors. A Likert scale was used for the responses where 1 = at every opportunity and 5 = never. The behaviors on the second part of the questionnaire were the same as those used in Part 1; however, they were listed in a different order. The Appendix gives the short form of the 14 behaviors surveyed by the questionnaire.

Based on the assumptions of LOC theory, an internal LOC individual is more likely to accept responsibility for his or her actions. The acceptance of responsibility for actions and the expected consequences of the outcomes of those actions suggests tthat those with an internal LOC will respond more ethically because of the believed behavior-outcome link. By contrast, an external LOC individual is less likely to accept responsibility for the consequences of unethical behaviors, since *externals* do not believe in the connection between their actions and the resulting outcomes. Without the assumed link between actions and expected consequences, *externals* are less likely to respond as ethically. Rotter's(1966) LOC instrument is used to designate student respondents as either internal or external. The instrument is a 29-item, double-statement survey developed to sample beliefs across various situations. As such, the LOC survey is considered to be a "generalized expectancy" measure. The LOC survey is designed to award a point every time that an *external*

answer is recorded by the subject. Since six of the LOC items are fillers and not counted in the scoring, the students' scores can range from zero to 23. Those students scoring less than or equal to10 on the LOC survey are designated as *internals*; those scoring greater than or equal to 11 were designated as *externals*for the analysis.

The SAS t-test procedures were used to analyze the differences in the students' mean responses on Parts 1 and 2 of the questionnaire. The five-point Likert scale mean responses for each behavior item were the dependent variable, and the LOC designation was used as the independent variable in the analyses.

Part 1 of the questionnaire requests the students to respond to how unethical they individually believe each of the described behaviors to be. The results of the t-test analysis for Part 1 are shown in Table 2. All 14 of the behaviors indicated significant differences at $\leq .05$ significance level. On all 14 items, the Internals responded with the greater ethical sensitivity. There is significant evidence that an internal LOC individual believes in the link between action and outcome even in situations involving additional individuals. This offers considerable support for the universal nature of LOC as a moderator of ethical beliefs and decisions and supports Hypothesis 1. Generally speaking, individuals who believe that they have some responsibility for the outcomes of their own behaviors responded with greater ethical sensitivity. These results support the belief that Internals are generally less susceptible to negative influence by others. The survey items were collaborative items, and follow the assumptions of the Internal LOC that one cannot accept attempts at influence by others because that is equivalent to assigning control to others.

Part 2 of the questionnaire used the same questionable behaviors and asked the students to consider them from a different perspective. The students were asked to indicate how often they believed most college students practiced each of the same collaborative or multiple▲individual behaviors. The surveyed behaviors were listed in a different order on Part 2 of the questionnaire; however, they have been converted to the corresponding Part 1 Item No. for ease of comparison. As predicted by Hypothesis 2, there were no significant differences in the mean responses by the LOC variable on 13 of the 14 survey items. Table 3 gives the statistical results for the significant item for Part 2 of the survey. A significant difference was found on only one item. The *external* students believe that the behavior described in Item No. 12, "Arranging with other students to give or receive answers by the use of signals," occurs more often in their environment than the *internal* students do. However, since both have high response means, neither believe that the behavior occurs very often. Since the LOC is considered to be a measure of the general control an individual is believed to have over his/her expected outcomes, differences on the LOC variable were not predicted and not generally indicated by the analysis. These results lend considerable support for Hypothesis 2.

Table 2 - Significant t-Test Results for Survey Part 1				
($\alpha \leq .05$)				
Model Item No.	Response Means*		t-statistic Results	p-value
	INTernal	EXTernal		
1	1.948	2.078	2.16	.0312
2	1.449	1.634	3.35	.0008
3	1.659	1.897	3.57	.0004
4	1.410	1.591	2.89	.0040
5	2.020	2.359	4.63	<.0001
6	2.701	3.021	3.73	.0002
7	1.630	1.986	5.25	<.0001
8	2.307	2.577	3.28	.0011
9	1.552	1.713	2.54	.0113
10	1.525	1.759	3.73	.0002
11	1.428	1.636	3.77	0002.
12	1.309	1.494	3.65	.0003
13	2.317	2.546	3.23	.0013
14	1.311	1.526	4.04	<.0001
On all of the 14 survey items, the INTernal students judged the questionable behaviors as more unethical than the EXTernal students did.				
** 1 = very unethical; 5 = not at all unethical*				

Table 3 - Significant t-Test Results for Survey Part 2				
($\alpha \leq .05$)				
Item No.	Model Response Means*		Results t-statistic	p-value
	INTernal	EXTernal		
	3.620	3.480	1.91	.0569
EXTernals believe the behavior occurs more often than INTernals				
* 1= at every opportunity; 5 = never				

DISCUSSION

This study examines differences in academic ethical responses of university students from the US . The survey that was used was set for the academic environment because it surveyed everyday ethical dilemmas familiar to the student subjects. The first part of the questionnaire surveyed students' ethical beliefs concerning the severity of the ethicality of the stated dilemmas. Based on the analysis by LOC, the *internal* LOC students exhibited greater ethical sensitivity on all of the 14 items. This predicted result is notable because of the collaborative nature of the surveyed dilemmas. That is, although the actions involve others in addition to the individual student, *internals* indicate greater responsibility for the action than do *externals*. This result suggests that LOC is such a strong personality descriptor that students with an internal LOC assume responsibility for their actions even in cases where they could easily transfer a part of the guilt to another person.

The second part of the survey queried the students' perceptions concerning how often they believe that the behaviors take place in their environment. Thirteen of the 14 behaviors revealed no significance differences on the LOC variable. The lack of significant differences indicates that LOC doesn't apply when assessing what is believed to be "in the environment." Some studies of students' ethical orientation or ethical decision making ask students about their own beliefs and actions and also about their perceptions of the beliefs and actions of their peers. To the extent that LOC applies differently to responses about one's own actions and the perceptions regarding the actions of others, the conclusions may not be valid. Therefore, this research offers evidence of the need to consider the moderating effect of LOC when designing ethical orientation surveys.

ACKNOWLEDGMENT

Financial support for this research was provided by Stephen F. Austin State University.

REFERENCES

Bass, K., Barnett, T., & Brown, G. (1999). "Individual Differences Variables, Ethical Judgments, and Ethical Behavior." *Journal of Business Ethics*, 9(2), 183-205.

Beu, D.S., Buckley, M.R., & Harvey, M.G. (2003). "Ethical Decision‐Making: A Multidimensional construct.", *Business Ethics: A European Review*, 12(1), 88-107.

Brownell, P. (1982). "A Field Study Examination of Budgetary Participating and Locus of Control." *The Accounting Review*, 56, 766-777.

Brownell, P. (1981). "Pariticpating in Budgeting, Locus of Control, and Organizational Effectiveness." *The Accounting Review*, 56, 844-860.

Chiu, R.K. (2003). "Ethical Judgment and Whistleblowing Intention: Examining the Moderating Role of Locus of Control." *Journal of Business Ethics*, 43, 65-74.

Geurin, V.T., & Kohut, G.F. (1989). "The Relationship of Locus of Control and Participative Decision Making among Managers and Business Students." *Mid-Atlantic Journal of Business*, 25, 57-66.

Hegarty, W.H., & Sims, H.P. (1979). "Organizational Philosophy, Policies, and Objectives Related to Unethical Decision Behavior: A Laboratory Experiment." *Journal of Applied Psychology*, 64, 331-338.

Hegarty, W.H., & Sims, H.P. (1978). "Some Determinants of Unethical Decision Behavior: An Experiment." *Journal of Applied Psychology*, 63, 451-457.

Jones, G.E., & Kavanaugh, M.J. (1996). "An Experimental Examination of the Effect of Individual and Situational Factors on Ethical Behavioral Intentions in the Workplace." *Journal of Business Ethics*, 15, 511-523.

Licata, M.P., Strawser, R.H., & Welker, R.B. (1966). "A Note on Participation in Budgeting and Locus of Control." *The Accounting Review*, 61, 112-117.

McCabe, D.L., Butterfield, K.D., & Trevino, L.K. (2004). "Academic Integrity: How widespread is Cheating and Plagiarism?" in *Restorative Justice on the College Campus: Promoting Student Growth, Responsibility, and Reawakening the Spirit of Campus Community*, D.R. Karp & T. Allena (eds.), Springfield, IL: Charles C. Thomas, 124-135.

McCuddy, M.K., & Peery, B.L. (1996). "Selected Individual Differences and Collegian' Ethical Beliefs." *Journal of Business Ethics*, 15, 261-272.

Newstrom, J.W., & Ruch, W.A. (1976). "The Ethics of Business Students: Preparation for a Career." *AACSB Bulletin*, 12 (April), 21-29.

Pratt, C.B., & McLaughlin, G.W. (1989). "An Analysis of Predictors of College Students' Ethical Inclinations." *Research in Higher Education*, 30, 195-219.

Rael, R. (2004). "Starting Over with Etics: How to Re-build Culture with Integrity." *Today's CPA*, Nov/Dec, 18-21.

Reiss, M.C., and Mitra, KI. (1998). "The Effects of Individual Differences Factors on the Acceptability of Ethical and Unethical Workplace Behaviors." *Journal of Business Ethics*, 17, 1581-1593.

Rogers, V., & Smith, A. (2001). "Ethics, Moral Development, and Accountants-in-training." *Teaching Business Ethics*, 5, 1-20.

Rotter, J.B. (1966). "Generalized Expectancies for Internal versus External Control of Reinforcement." *Psychological Monographs*, 80 (Whole No. 609).

Smith, A., Coates, J.D., & Deis, D.R. (1999). "Are Ethical Responses Linked to Locus of Control?" *Teaching Business Ethics*, 2(3), 249-260.

Spector, P.E., & Michaels, C.E. (1986). "Personality and Employee Withdrawal: Effects of Locus of Control on Turnover." *Psychological Reports*, 59, 63-66.

Stevens, G.E. (1984). "Ethical Inclinations of Tomorrow's Citizens: Actions Speak Louder?" *Journal of Business Education*, 59, 147-152.

Storms, P.L., & Spector, P.E. (1987). "Relationships of Organizational Frustration with Reported Behavioural Reactions: The Moderation Effect of Locus of Control." *Journal of Occupational Psychology*, 60, 227-234.

Terpstra, E.E., Ryes, G.C., & Boker, D.W. (1991). "Predictors of Ethical Decisions Regarding Insider Trading." *Journal of Business Ethics*, 10, 699-710.

Trevino, L.K., & Youngblood , S.A. (1990). "Bad Apples in Bad Barrels: A Causal Analysis of Ethical Decision▲ making Behavior." *Journal of Applied Psychology*, 75, 378-385.

Tyson, T. (1990). "Believing that Everyone Else is Less Ethical: Implications for Work Behavior and Ethics Instruction." *Journal of Business Ethics*, 9, 715-721.

APPENDIX - Short Form of Behaviors Surveyed

No.	Behavior
1	Citing someone else's work as your own
2	Failure to report unfavorable grading errors
3	Copying homework and turning it in as your own
4	Using cheat sheets during an exam
5	Not contributing your fair share of a group project
6	Falsifying or fabricating a bibliography/references list
7	Studying from someone else's notes
8	Visiting a professor after an exam, attempting to bias grading
9	Obtaining an old exam from a previous semester or quarter
10	Changing a test paper from the original handed in
11	Making improper use of another's computer file/program

THE PERSPECTIVE OF FACULTY HIRED AFTER AACSB ACCREDITATION ON ACCREDITATION'S IMPACT AND IMPORTANCE

Wayne A. Roberts, Jr., Southern Utah University
Roy Johnson, Southern Utah University
John Groesbeck, Southern Utah University

ABSTRACT

Obtaining AACSB accreditation is a long, resource consuming exercise. In this study 62 faculty hired by 24 schools that had recently received AACSB accreditation between 1997 and 2001 were surveyed to determine their impressions of the impact and importance of accreditation on various stakeholders. Their responses are subsequently compared to the responses of faculty who were employed by their school prior to gaining AACSB accreditation. Overall, accreditation was perceived as being beneficial to the business school, students, and faculty, and to the employers of students. This is consistent with the perception of faculty who were employed at newly accredited institutions prior to receiving AACSB accreditation. Newly hired faculty perceive that they value research and teaching more than established faculty. In contrast, previously published results indicate faculty employed prior to receiving accreditation believe those hired since accreditation value teaching less. Importantly, everything else being equal, respondents decidedly prefer working at AACSB accredited schools.

INTRODUCTION

Obtaining AACSB (Association to Advance Collegiate Schools of Business) accreditation is a major undertaking. It takes time, diverts a lot of administrative and faculty time from other activities, is fraught with uncertainty, and takes money. A fundamental question is whether or not it is worth the effort and expense.

A previous paper addressed the results of a survey of faculty who went through the accreditation process regarding their perceptions of the impact of AACSB accreditation (Roberts, Johnson & Groesbeck, 2004). Faculty from recently accredited schools rated the impact of accreditation on the business school, faculty who were with the school before accreditation, faculty hired since accreditation, the programs, and students and employers of students.

In this paper we seek to supplement the insights gained from that study by focusing on faculty who were newly hired by schools that had recently received AACSB accreditation.

Understanding the perceptions of potential new faculty should be an important consideration in deciding whether or not to pursue AACSB accreditation. How do new hires differ from established faculty with regard to the perceived impact of AACSB accreditation? Do those who are drawn to newly accredited programs feel differently about accreditation than those who joined programs that were not then accredited, but who later instigated the accreditation process? Are the values of new faculty different from established faculty, and is AACSB accreditation an important consideration in judging alternative schools?

AACSB ACCREDITATION

There are currently 466 AACSB accredited business programs (AACSB International, 2004). The cost of gaining this accreditation can be high. The direct costs -- application fees, conference fees, air fares, meals, and the ever-present cost of hiring consultants -- can be well over $50,000. If the school is aggressive, these costs typically approach $100,000 (see Roberts, Johnson & Groesbeck, 2004).

In addition to the direct costs of pursuing accreditation is the time and effort that faculty and administrators are required to invest to achieve compliance. While it is true that a substantial amount of the effort required by AACSB is really nothing more than the good management practices that most schools should be doing whether or not they are pursuing accreditation, there remains a large amount of reporting and compliance work, which crowds out more important work. Overworked faculty do most of it (Holmes, 2001), and they see such work as unrewarded service work (Henninger, 1998).

With regard to ongoing costs, inevitably a school in candidacy will have to hire academically qualified faculty, and those faculty are not cheap. For the sake of comparison, the average salary of an associate professor at a public AACSB accredited school was about $85,000 in 2002-03. The average salary of an associate professor at a non-AACSB accredited school was about $69,000 for that same year. Hiring a new associate professor with AACSB-appropriate credentials to anchor an academic program could easily cost more than $91,000 in that same year, depending on the discipline (see AACSB International, 2002). Salary gaps between existing business faculty and newly hired academicians can be very large, and this can cause problems with existing faculty, and with university administrators. Not insignificantly, faculty in other disciplines outside of business, who do not like the salary gap as it is, may become even more upset when market salaries for new AACSB-appropriate faculty starts to take place.

One of the most visible consequences of AACSB accreditation has been an increased focus on research. In fact, Udell, et al. (1995) note that, "Discussions of the validity and desire for AACSB accreditation generally become discussions of the seeming dichotomies of teaching and research" (p. 108). They found that faculty of AACSB accredited schools published significantly more journal articles than faculty of institutions denied accreditation, though there was no difference

in broadly defined "scholarly activities." Ehie and Karanthanos (1994) found that, while overall emphasis on teaching had grown, "accredited institutions perceive instructional responsibilities as less important and intellectual contributions as more important than do nonaccredited institutions."

This focus on research by the AACSB has been downplayed by new "mission driven" standards implemented in 1994. Several studies have examined the impact of these new standards, especially as they concern research requirements. Henninger (1998) found "only modest changes in faculty selection and work resulting from the new standards." Similarly, Jantzen (2000) found that, "The adoption of 'mission-related' standards, by itself, has not resulted in a change in either the number or character of schools being accredited." Arlinghaus (2002), in a study of AACSB accounting programs, found that "the expectation for the volume of publication has increased at the majority of respondent institutions for both tenured and untenured faculty."

AACSB has again been revising these standards through a blue ribbon commission over the past few years, and formally adopted new standards in 2003. These new standards will take effect very soon, and it is too early to tell what will happen to the entire "mission driven" focus of the organization. Some changes have taken place recently that hint a return to more of a traditional research-based focus within AACSB. These changes include dismantling the Candidacy Committee as a separate group within the organization, which appeared to be one of the strongest advocates for the "mission-driven" movement. From now on, all accreditation processes for both new and existing schools will be managed by the Accreditation Committee, which tends to be populated more by the larger and older business schools which were the traditional accredited members of AACSB. Further, in about 2002 AACSB began requiring that accredited and candidacy schools annually report peer-reviewed journal articles by each faculty member over the previous five years, regardless of the school's mission. For good or bad, the emphasis on research appears to remain.

It thus seems clear that AACSB accreditation has a major impact on business schools. Roberts, Johnson & Groesbeck surveyed faculty who were employed by recently accredited schools prior to AACSB accreditation to determine their impressions of the impact of accreditation on various stakeholders (2004). Overall, accreditation was perceived as being beneficial to the business school, students, and faculty hired since accreditation, and to the employers of students. These faculty members, however, did not perceive accreditation as helping them personally.

Given that accreditation may result in a cultural change and a resetting of priorities and rewards in a school, it was thought that more insight into the value of accreditation might be gained by contrasting the perspectives of faculty who are hired after accreditation with those of faculty hired prior to accreditation. New faculty may use accreditation as an indication of whether schools meet their career expectations. They have actively chosen a school with the priorities that accreditation gives them. Faculty hired prior to accreditation, on the other hand, may well have had different expectations when they joined a program and have seen the established culture changed. Many of these faculty members may be moving grudgingly in a direction they recognize as important to their school's future (Roberts, Johnson & Groesbeck, 2004).

The purpose of this study is to assess the impact of AACSB accreditation on the schools through the eyes of faculty hired since accreditation, to assess the importance of accreditation on their decision to accept employment at the institution, and to compare the impressions of these faculty members to those who experienced the institutions prior to accreditation as well as subsequent to it.

DATA AND SAMPLE COLLECTION METHOD

A list of business schools which obtained AACSB accreditation during the years 1997 through 2001 was obtained, and six United States business schools were randomly chosen for each year, resulting in a total of 30 schools. Table 1 lists the chosen schools. Email addresses for business school employees were obtained via the Internet, except for one university, which agreed to forward emails and questionnaires to their faculty.

Each email address was sent a cover letter with links to questionnaires on two different occasions. Ten schools were emailed April 12 and April 29, 2001, and the other 20 schools were surveyed November 21, 2001 and January 28, 2002. Faulty who were with the institution prior to receiving AACSB accreditation were requested to access one questionnaire, the results of which are discussed elsewhere (Roberts, Johnson & Groesbeck, 2004). Non-faculty were asked to reply to the email and indicate that they were not faculty. Faculty hired since accreditation, the focus of this study, were asked to fill out a questionnaire designed for them.

Eliminating those that indicated they were not faculty resulted in a total of 1121 email addresses. Sixty-two respondents indicated they were faculty hired since accreditation, while 221 respondents were hired prior to accreditation. Hence, the minimum response rate was 25.25%. Depending on the number of non-faculty that remained on the list, the response rate might be considerably higher. It is not, of course, known how the response rate of those hired since accreditation compares to the response rate of those hired prior to accreditation.

Table 1: Recently AACSB accredited schools selected for study

St. Mary's University	The U. of Tampa
Marshall University	Illinois Institute of Tech.
SW Texas State	Chapman University
Fairfield University	Seattle Pacific U.
New Jersey Institute of Tech	U. of Mass. – Dartmouth
Pace University	Iona College
Jacksonville State U.	Niagara U.
Henderson State U.	Winston-Salem State U.
Rice University	No. Carolina St. U.

Table 1: Recently AACSB accredited schools selected for study	
Longwood College	Indiana University Kokomo
Coastal Carolina U.	Fairleigh Dickinson U.
The College of NJ	U. of Mass. – Boston
Quinnipiac University	Michigan Tech. U.
Truman State	Long Island U.
St. Joseph's U.	No. Dakota State University

QUESTIONNAIRE

The questionnaire consisted of 17 Likert statements and 6 questions regarding the status and history of the respondent. It was developed using Microsoft Frontpage, and submitted questionnaires were automatically dumped into an Excel file, thus eliminating input error. For the purposes of analysis, Likert responses were coded as follows: -2 = strongly disagree, -1=disagree, 0=neither agree nor disagree, 1=agree, and 2=strongly agree. For some statements respondents were given a 'not applicable' choice, which was treated separately.

RESPONDENT CHARACTERISTICS

Table 2 summarizes data regarding the respondents. With regard to rank, 56.5% were assistant professors, 14.5% were associate professors, and 12.9% were full professors. With regard to tenure, at least 14.5% had tenure. With regard to discipline represented, 16.1% were from accounting, 11.3% taught economics, 8.1% finance, 4.8% information systems, 21.0% management, 19.4% marketing, and 8.1% quantitative methods. With regard to their academic experience, 40% reported that their current position was their first faculty position, while 30% reported that they had more than seven years experience at other academic institutions.

Based on these results there is little reason to suspect that the respondents do not constitute a representative sample of faculty hired at institutions subsequent to AACSB accreditation.

RESULTS

Responses to the survey are provided in table 3. The exact wording of the Likert statements is provided. For presentation purposes the order of the statements have been rearranged, and do not reflect the sequence in which they appeared on the questionnaire. The main discrepancy is that overall assessment questions were asked near the end of the questionnaire, just before demographic data was collected. The 2-tailed significance value provided reflects the probability that you would

get the sample mean if the null hypothesis, that the mean value equals 0, is true. Zero represents the neutral point (neither agree nor disagree).

Table 2: Selected respondent characteristics

Number of Respondents by Year Accredited		Areas of Teaching Responsibility	
Year	Number	Discipline	Number
1997	19	Accounting	10
1998	20	Economics	7
1999	13	Finance	5
2000	5	Information systems	3
2001	5	Management	13
TOTAL	62	Marketing	12
		Quantitative methods	5
		Other response	7
		TOTAL	62
Respondent Tenure Status			
Tenured	9		
Not tenured	48		
Missing			
Other	5		
TOTAL	62		

Academic Experience			Faculty Rank	
	Other schools	Current school		
1 year or less	3	18	Instructor	6
More than 1 to 3	5	29	Assistant Professor	35
More than 3 to 5	4	8	Associate Professor	9
More than 5 to 7	6	1	Full Professor	8
More than 7	18	4	Other	3
1st faculty position	24	-	Missing	1
Missing	2	2	TOTAL	62
TOTAL	62	62		

Table 3: Perceived AACSB accreditation impacts							
Likert Statement (n)	Mean (Standard deviation)[a]	Sig.[b] (2-tailed)	SA[a] (+2)	A (+1)	N (0)	D (-1)	SD (-2)
On faculty and faculty choice							
Everything else being equal, I would prefer to work at an AACSB accredited institution. (61)	1.62 (.71)	.000	73.8	16.4	8.2	1.6	0.0
Overall, I believe AACSB accreditation has benefitted faculty who were here prior to accreditation. (60)	.57 (1.02)	.000	18.3	38.3	26.7	15.0	1.7
Overall AACSB accreditation benefits me (61)	1.18 (.90)	.000	41.0	44.3	8.2	4.9	1.6
Overall, AACSB accreditation benefits new faculty (60)	1.18 (.89)	.000	43.3	36.7	16.7	1.7	1.7
On business school							
Overall, AACSB is good for the business school (61)	1.41 (.64)	.000	47.5%	47.5%	3.3%	1.6%	0.0%
AACSB Accreditation helps our business school compete for…							
…financial resources (62)	.90 (.90)	.000	29.0	38.7	25.8	6.5	0.0
…students (62)	1.05 (.82)	.000	30.6	48.4	16.1	4.8	0.0
…What I consider to be appropriate faculty (61)	1.18 (1.01)	.000	47.5	34.4	8.2	8.2	1.6
AACSB accreditation process helps ensure that we have, and will continue to have, a quality program (61)	.90 (.93)	.000	24.6	50.8	18.0	3.3	3.3
On new versus former faculty Compared to faculty who worked here prior to AACSB accreditation efforts, new faculty...							
… generally value research more. (57)	1.16 (.73)	.000	33.3	50.9	14.0	1.8	0.0
…generally value teaching more. (57)	.40 (.92)	.002	15.8	21.1	52.6	8.8	1.8
…generally value university/public service less. (56)	-.18 (.834)	.115	3.6	12.5	50.0	30.4	3.6
…have better contracts. (51)	.27 (1.00)	.056	7.8	37.3	35.3	13.7	5.9

Table 3: Perceived AACSB accreditation impacts

Likert Statement (n)	Mean (Standard deviation)[a]	Sig.[b] (2-tailed)	SA[a] (+2)	A (+1)	N (0)	D (-1)	SD (-2)
On students and employers							
Overall, I believe AACSB accreditation benefits students. (61)	1.11 (.84)	.000	32.8	52.5	9.8	3.3	1.6
Overall, I believe AACSB accreditation benefits employers of our students. (61)	.62 (.76)	.000	11.5	44.3	39.3	4.9	0.0
Overall assessment and recommendation							
Overall, I believe AACSB accreditation is worth the effort to obtain it. (61)	.97 (.91)	.000	31.1	41.0	23.0	3.3	1.6
Overall, I believe AACSB accreditation is something I would recommend to other schools. (61)	.93 (.87)	.000	26.2	47.5	21.3	3.3	1.6

[a] 5-point scale where +2=strongly agree (SA), +1=agree (A), 0= neither agree nor disagree (N), -1=disagree (D), and –2 =strongly disagree (SD).
[b] 2-tailed significance associated with H:mean=0

Perhaps one of the most interesting results presented in table 3 is represented by the first item regarding faculty and faculty choice impacts. As shown, the mean score for the statement "Everything else being equal, I would prefer to work at an AACSB accredited institution" was statistically significant with a mean value of 1.62. This mean was the highest of the survey, and more respondents strongly agreed with that statement than with any other. Fully 90.2% agreed or strongly agreed with the statement, while only 1.6% of the respondents disagreed, and none strongly disagreed. Among the 24 respondents in their first academic position, 19 (79.2%) strongly agreed, 4 (16.7%) agreed, and only 1 (4.2%) neither agreed nor disagreed with the statement. Clearly, at least for faculty hired at AACSB accredited institutions, such accreditation is very important. If we make the assumption that among academically qualified faculty for whom AACSB accreditation is not important there is no bias against accepting employment at accredited schools, it appears that AACSB accreditation is a critically important characteristic among academic job seekers.

The other items regarding faculty and faculty choice show that respondents perceive AACSB accreditation as benefiting them, new faculty, and faculty who were there prior to accreditation. Over 85% strongly agreed or agreed with the statement regarding AACSB accreditation benefiting them, 80% strongly agreed or agreed with the statement regarding new faculty, and almost 57% agreed with the statement regarding faculty hired prior to receiving accreditation.

With regard to the impact of accreditation on the business school, respondents, in general, agree with the statement that it has been positive. As shown in table 3, the mean response for the

overall statement, 1.41, was significantly above zero. Ninety-five percent agreed or strongly agreed that it was good for the business school, and only 1.6% disagreed with the statement. The majority of respondents agree that AACSB accreditation helps the business school compete for financial resources, students, and faculty. Further, 75.4% agreed or strongly agreed that accreditation helps ensure that they have, and will continue to have, a quality program. Only 6.6% disagreed with that statement.

With regard to statements concerning values of faculty hired since accreditation as compared to faculty employed there prior to accreditation, respondents saw the newer faculty as valuing research and teaching more. Mean scores for the statements concerning research and teaching, provided in table 3, were both positive and statistically significant. Importantly, though, 52.6% neither agreed nor disagreed that new faculty value teaching more, and an additional 10.6% disagreed or strongly disagreed with the statement. With regard to the statement that new faculty have better contracts, while the mean value was positive the level of significance of the univariate test was.056, which, it could be argued, is marginally significant. Over 35% neither agreed nor disagreed with that statement, and almost 20% disagreed or strongly disagreed with it. Results regarding the university/public service component of faculty work were not statistically significant, and 50% neither agreed nor disagreed with that statement.

Two of the most important groups to any university are, of course, students and employers of students. As shown in table 3, respondents agreed, as a group, that AACSB accreditation benefits both. With regard to the statement that AACSB accreditation benefits students, the mean response was 1.11, and was statistically significantly different from zero, with 72.1% agreeing or strongly agreeing. Only 9.8% neither agreed nor disagreed, and only 4.9% disagreed or strongly disagreed with the statement. With regard to the statement that AACSB accreditation benefits the employers of students, the mean response was .62, which was statistically significant, with 55.8% agreeing or strongly agreeing, 39.3% neither agreeing nor disagreeing, and only 4.9% disagreeing or strongly disagreeing.

At the bottom of table 3 are the responses to two statements intended to assess respondents' general feelings about the value of accreditation. Over 72% agreed or strongly agreed that AACSB is worth the effort to obtain it, while only 4.9% disagreed or strongly disagreed. The mean response was .97 and statistically significant. Over 73% agreed or strongly agreed with the statement that they would recommend it to other schools, while only 4.9% disagreed or strongly disagreed. Again, the mean response was .93 and statistically significant.

It is interesting to compare how perceptions of faculty hired since accreditation compare to perceptions of faculty who were present prior to their school's gaining AACSB accreditation. Table 4 compares the mean responses from the two groups, and tests the hypotheses that the means are equal, without assuming that the variances are equal. The mean responses from new faculty are taken from table 3, while responses from faculty who went through the accreditation process are from Roberts, Johnson & Groesbeck (2004). As mentioned earlier, the two surveys were

administered at the same time to faculty at the same institutions. It should be noted, however, that the Likert statements were worded slightly differently in the two surveys, in an attempt to make them more meaningful to the respondents. Basically, the statements provided established faculty used the past tense, while those to new faculty used the present tense. For example, whereas new faculty were asked to indicate their agreement with "Overall, I believe AACSB accreditation is good for the business schools," the wording for established faculty was "Overall, AACSB accreditation has been good for the business school." The exact wording for new faculty is provided in table 3, and the exact wording for established faculty is provided in Roberts, Johnson & Groesbeck (2004). Nevertheless, it is believed that the comparisons are appropriate and meaningful.

Restricting the discussion to those statements where the achieved level of significance was less than .01, new faculty tended to agree more that accreditation helped the respondents themselves, as well as new and established faculty in general. Further, they agreed more that AACSB accreditation helps the business school, overall, and more specifically when competing for appropriate faculty. Perhaps surprisingly, new faculty tended to agree that new faculty value teaching more than established faculty, while established faculty tended to disagree. Finally, new faculty agreed more that AACSB accreditation benefits students and employers.

Expanding the discussion to include achieved levels of significance between .01 and .05, the results suggest that there may be significant disagreement between the two groups regarding university/public service, and on the extent to which the two groups recommend pursing AACSB accreditation to other schools. With regard to university/public service, new faculty tended to disagree with the statement suggesting new faculty value it less than established faculty, while the mean response from established faculty was above zero, although not statistically significantly so. Importantly, the modal response for both groups was the neither agree nor disagree category. With regard to recommending accreditation to other schools, the new faculty are more inclined to do so.

Table 4: Evaluation of differences between new and established faculty[a]

Likert Statement Subject	Mean Values[b] (n)				
	New Faculty	Established Faculty	Difference	t	Sig. (2-tailed)
Impact on faculty					
Overall, accreditation has benefited me	1.18 (61)	.13 (216)	1.055	7.524	.000
Overall, accreditation has benefited faculty who were here prior to accreditation	.57 (60)	-.01 (220)	.580	3.909	.000
Overall, accreditation's benefited new faculty	1.18 (60)	.79 (214)	.389	2.985	.004

Table 4: Evaluation of differences between new and established faculty[a]

Likert Statement Subject	Mean Values[b] (n)				
	New Faculty	Established Faculty	Difference	t	Sig. (2-tailed)
Impact on business school					
Overall, accreditation has been good for the business school	1.41 (61)	1.06 (221)	.346	3.344	.001
Accreditation helps compete for financial resources	.90 (62)	.73 (219)	.177	1.321	.189
Accreditation helps compete for students	1.05 (62)	.83 (218)	.223	1.838	.069
Accreditation helps compete for appropriate faculty	1.18 (61)	.77 (217)	.406	2.793	.006
Accreditation helps ensure program quality	.90 (61)	.72 (220	.179	1.280	.203
On new versus established faculty					
New faculty value research more	1.16 (57)	1.04 (217)	.121	1.081	.283
New faculty value teaching more	.40 (57)	-.32 (217)	.726	5.194	.000
New faculty value university/public service less	-.18 (56)	.09 (215)	-.272	-2.101	.038
New faculty have better contracts	.27 (51)	.48 (211)	-.204	-1.286	.202
Impact on students, employers, and overall assessment					
Overall, accreditation has benefitted students	1.11 (61)	.60 (220)	.510	4.010	.000
Overall, accreditation has benefitted employers	.62 (61)	.24 (219)	.386	3.384	.001
Overall, accreditation's worth the effort	.97 (61)	.78 (221)	.184	1.341	.183
Overall, I'd recommend accreditation to other schools	.93 (61)	.66 (220)	.271	2.030	.045

[a] 2-tailed significance associated with H: difference=0, equal variances not assumed.
[b] 5-point scales where +2=strongly agree, +1=agree, 0= neither agree nor disagree , -1=disagree, and -2 =strongly disagree

DISCUSSION OF RESULTS

The results suggest that faculty hired at AACSB accredited institutions view such accreditation very positively. Their perception is that it helps the business school compete for students, faculty, and financial resources, and helps ensure a quality program. Their perception is that it benefits all business faculty, and in particular, themselves. Perhaps because of this, they prefer to work at AACSB accredited institutions.

New faculty also believe that AACSB accreditation benefits students and their future employers, believe AACSB accreditation is worth the effort, and would recommend AACSB accreditation to other schools.

With regard to comparing themselves to faculty established at their institution prior to AACSB accreditation, new faculty see themselves as valuing research and teaching more. With regard to whether or not they have better contracts, while they tended to agree, the significance level of the test that they neither agree nor disagree was only .056.

Comparing established faculty responses to new faculty hires, the primary difference seems to be that newer faculty more strongly agree that AACSB accreditation benefits faculty, helps their school compete for appropriate faculty, helps the business school overall, and benefits students and employers. Both groups perceive that AACSB accreditation changes the values of the organization, in that new faculty value research more. There is disagreement between the two groups regarding teaching: new faculty think they value teaching more, while established faculty believe new faculty value teaching less.

CONCLUSIONS AND COMMENTS

While the sample size obtained for this study was relatively small, statistically significant and practical results were obtained. A limitation of this study concerns not the raw numbers, but the unknown response rate, and the possibility that the sample was heavily weighted towards faculty hires familiar with AACSB accreditation. That is, newly minted Ph.D.s unfamiliar with AACSB International would not have a basis for evaluating its significance, and hence may have been less inclined to respond to this survey.

With this limitation in mind, it is still possible to state that among those familiar with AACSB International (at least those that end up in AACSB accredited institutions), AACSB accreditation is an important institutional characteristic. Further, such faculty, like those who lived through the accreditation process, believes it helps the business school, the faculty, students and employers, enough so that they would recommend it to others.

One perceived consequence of AACSB accreditation is that the character of the faculty changes in at least one respect: New hires value research more. It is not clear whether or not this means they also value teaching less: faculty hired after accreditation do not believe so, but as

reported in Roberts, Johnson & Groesbeck (2004), established faculty do. An interesting question is whether discrepancies of perceptions sets up conditions that generate conflict.

For those contemplating pursuing AACSB accreditation, the majority of faculty that witnessed changes at schools that successfully went through the process, as well as faculty hired since accreditation, believe its achievement is worth the effort. Further, at least among those that end up at accredited institutions and know what it stands for, it tends to be an important institutional characteristic that presumably impacts their choice of employer.

REFERENCES

AACSB International (2002). *2001-2003 Salary Survey Report*. St. Louis: Knowledge Services.

AACSB International (n.d.). *Why Join AACSB?* Retrieved July 15, 2004, from http://www.aacsb.edu/members/why.asp

Arlinghaus, B. P. (2002). The environment for professional interaction and relevant practical experience in AACSB-accredited accounting programs. *Journal of Education for Business*, 77 (4), 38-45.

Ehie, I. C. & D. Karathanos (1994). Business faculty performance evaluation based on the new AACSB accreditation standards. *Journal of Education for Business*, 69 (5), 257-263.

Jantzen, R. H. (2000). AACSB mission-linked standards: effects on the accreditation process. *Journal of Education for Business*, 75 (6), 343-48.

Henninger, E. A. (1998). Dean's role in change: the case of professional accreditation reform of American collegiate business education. *Journal of Higher Education Policy & Management*, 20 (2) 203-214.

Henninger, E. A. (1998). Perceptions of the impact of the new AACSB standards on faculty qualifications. *Journal of Organizational Change Management*, 11 (5), 407-424.

Holmes, L. (2001). Seeking accreditation is well worth the work. *Black Issues in Higher Education,* 18 (20), 152-153.

Roberts, W. A. Jr., R. Johnson and J. Groesbeck (2004). The faculty perspective on the impact of AACSB accreditation. *Academy of Educational Leadership Journal*, 8 (1), 111-125.

Udell, G.G., S. Parker, and C. Pettijohn (1995). An examination of the research productivity of marketing faculty at newly minted AACSB schools. *Journal of Marketing Theory and Practice,* 3 (2), 106-113.

CONFIRMATORY FACTOR ANALYSIS OF THE PRINCIPAL SELF-EFFICACY SURVEY (PSES)

R. Wade Smith, Louisiana State University
A. J. Guarino, Auburn University

ABSTRACT

This article describes the development and constructs validity of the Principal Self-Efficacy Survey (PSES). The item selection was based on the theoretical framework proposed by Bandura. Fourteen-items assessing two factors Instructional Leadership (nine items) and Management Skills (five items) and a demographic questionnaire comprised the PSES. Items were scored on a 1 to 4 Likert-type scale. Participants were two hundred eighty-four principals. Construct validity was supported by confirmatory factor analysis using AMOS 5.0. In conclusion, the PSES provides a promising measure of principal perceptions of their ability to effectively function in the areas of instructional leadership and management.

INTRODUCTION

Bandura (1997) defines self-efficacy as: ". . . beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (p.3). According to Bandura, self efficacy influences, (1) the courses of action people choose to pursue, (2) how much effort people will put forth in a given endeavor, (3) how long they will persevere in the face of obstacles and failure, (4) people's resilience to adversity, (5) whether someone's thought patterns are self-hindering or self-aiding and (6) how much stress and depression is experienced in coping with taxing environmental demands.

The central role of self-efficacy in human agency makes it an important and useful construct for empirical research. Because self-efficacy is a task-specific construct (Bandura, 1997), any attempt to measure self-efficacy should be contextually sensitive to the setting in which the behaviors occur. A rich and robust body of literature documents the relationships between self-efficacy beliefs for teachers and students and their relationship to teaching and learning (e.g., Pajares, 1996; Tschannen-Moran, Hoy, and Hoy, 1998). However, a literature search for journal articles on *principal self-efficacy and instructional effectiveness* produced no articles specific to the topic. Currently there is tremendous interest in the role of the principal in affecting substantive, long-term improvement in schools. For example, the federal government, in The No Child Left Behind Act has weighed in with a mandate that principals in poorly performing schools shall be replaced if improvement is not forthcoming.

Given the central role that principals are expected to perform in maintaining quality teaching and learning environments in schools, it is important to begin to conceptualize and operationalize measures of principal self-efficacy. The following sections detail the development of the Principal Self-Efficacy Survey (PSES) along with its attendant psychometric properties.

ITEM GENERATION

The generation of items for the PSES used the rational-empirical approach to instrument development (Burisch, 1984). The rational component drew upon the knowledge and experience of professionals working as principals and the research literature to suggest potential items. The empirical component selected or rejected items based on their psychometric properties. The scale configuration was based on the theoretical framework proposed by Bandura. Fourteen-items assessing two factors Instructional Leadership (nine items) and Management Skills (five items) and a demographic questionnaire comprised the PSES. Items were scored on a 1 to 4 Likert-type scale.

ITEM SELECTION

The 14 items were then checked for violations of normalcy through the SPSS Statistical Package Version 11.0 (SPSS Inc., 2001), explore function. Items would be considered for elimination if they had a skew value equal or greater than two and kurtosis value equal or greater than seven.

PARTICIPANTS

Two hundred and eighty-four principals returned completed and valid surveys representing twelve states (5 in the southeast, 2 in the Midwest, 2 in the west, 2 in the northeast, and Alaska). There are 74 elementary schools, 30 middle schools, and 31 high schools represented in this study. Sixty-six percent of the respondents are males. Ethnic representation included 83% white, 14% black, and 1.4% other. Nearly 47% of the respondents indicated that they have a master's degree plus 30 hours and approximately 10% of respondents have an earned doctorate. The majority of the responses (54%) came from rural schools, while 17% were from suburban schools and 25% were from urban schools

RESULTS

Because missing data appeared to be randomly scattered among the variables, a full information maximum likelihood (FIML) imputation was performed to estimate missing data. The factor structures were examined using a confirmatory factor analysis. A series of models were tested

in the following order: (a) a single-factor *g* model in which all items were free to load on only one common factor; (b) an orthogonal two-factor model in which each factor was set to be independent of each other; (c) a correlated two-factor model in which the factors were to each other. The first two models were included to aid in the assessment of the correlated two- factor model.

The models were examined by AMOS version (5.0) maximum likelihood factor analysis (Arbuckle, 2004). The models were evaluated by a variety of fit measures that are classified as absolute, relative, parsimonious, and population discrepancy. Absolute fit measures assess how well the proposed interrelationships among the variables match the interrelationships among the actual interrelationships. The measure of absolute fit used in this study was the chi-square test because AMOS does not provide other absolute measures when missing data is estimated with the FIML imputation procedure. Measures of relative fit compare the hypothesized model to the null model. The relative fit measures employed in this study were the Comparative Fit Index (CFI) (Bentler, 1990), the Tucker-Lewis Index (TLI) (Bentler and Bonett, 1980). Measures of parsimonious fit attempt to determine if the overall fit of the model has been accomplished by overfitting the data. The parsimonious fit measure in this study was the chi-square divided by the degrees of freedom. Lastly, population discrepancy measures are estimates from the sample coefficients to the population coefficients. The population discrepancy measure in this study was the Root Mean Square Error of Approximation (RMSEA) (Browne and Cudeck, 1993). Models were compared by examining differences in values of chi-square to identify statistically significant variations among the models. The fit indices for the three models are presented in Table 1.

Table 1: Fit Indices for the Three Models

Factor Model	χ^2	df	χ^2 / df	CFI	TLI	RMSEA
Single (*g*)	180.37*	77	2.34	.993	.991	.069
Orthogonal	218.60*	77	2.84	.991	.987	.081
Correlated	127.1*	76	1.67	.997	.995	.049

* $p < .05$.

The chi-square test for differences revealed that the correlated two-factor model is superior to the other models. The correlated two-factor model yielded acceptably high goodness of fit indices (i.e., > .99) for both the CFI and the TLI. The RMSEA achieved a value of .049 indicating a close fit between the sample coefficients and the estimated population coefficients. The correlation between the two factors is .69 demonstrating discriminate validity.

The factor loadings are provided in Table 2. All items loaded statistically significantly ($p < .01$) and demonstrated practical significance with loadings greater than .40 on their respective factors.

Table 2. Item factor loadings		
Factor Loadings of the Principal Efficacy Survey		
Item	Instructional Leadership	Management Skills
Q1	.69	
Q2	.62	
Q3	.59	
Q4	.65	
Q5	.66	
Q6	.64	
Q7	.59	
Q8	.65	
Q9	.61	
Q10		.66
Q11		.77
Q12		.47
Q13		.58
Q14		.44

CONCLUSION

This study provides empirical evidence that the PSES operationalizes the latent constructs of instructional leadership and management skills for principals. Individual items demonstrated construct validity, (i.e., the items were shown to measure their respective hypothetical construct and factor loadings were all significant, $p < .01$). The instructional leadership and management constructs are both considered essential to principal effectiveness and as such, the PSES provides a promising measure for furthering understanding of self-beliefs of principals.

Because this research was exploratory in nature, further research is suggested to replicate the initial results. Also, future research should attempt to determine if the factor structure holds for various levels of the principalship (i.e., elementary, middle, and high school). Future research incorporating other important elements of principal self-efficacy beliefs (e.g., conflict resolution) would also be suggested. Finally, it would be important to understand principal self-efficacy for instructional effectiveness within the broader context of constructs known to be important for creating and facilitating an effective learning environment in schools. With this in mind, future studies should investigate the relationships between principal self-efficacy and other important constructs such as school culture, teacher self-efficacy, and student self-efficacy.

REFERENCES

Arbuckle, J.L. (1999). *Amos 4.0 User's Guide*. Chicago: Small Waters Corporation.

Bandura, A. (1997). *Self-efficacy: the exercise of control.* New York: Freeman.

Bandura, A. (2001). *Guide for constructing self-efficacy scales (Revised).* Available from Frank Pajares, Emory University.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.

Bentler, P.M. and Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588-606.

Browne, M.W. and Cudeck, R. (1993). Alternative ways of assessing model fit. In Bollen, K.A. and Long, J.S. [Eds.] *Testing structural equation models.* Newbury Park, California: Sage, 136-62.

Burisch, M. (1984). Approaches to personality inventory construction: a comparison of merits. *American Psychologist, 39*(3) 214-227.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 4* (66) 4. 543 – 578.

Tschannen-Moran, M.., Hoy, A.W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research. 68*(2) 202 – 248.

APPENDIX A
PRINCIPAL SELF-EFFICACY SURVEY
PRINCIPAL SURVEY

INSTRUCTIONS

This administrator survey asks you to make a series of judgments about your experiences as a head administrator for a school. You are asked to read the following items and rate the strength of your beliefs in your abilities to attain the following outcomes. These items should be answered from your perspective as a school principal working to produce an effective teaching and learning environment. You are to indicate the degree to which you agree or disagree with each statement by darkening the appropriate oval.

Scale 1=Very Weak Beliefs in My Abilities (VW)
2=Weak Beliefs in My Abilities (W)
3=Strong Beliefs in My Abilities (S)
4=Very Strong Beliefs in My Abilities (VS)

STATEMENTS:

My beliefs in my abilities to…

1. influence teachers to utilize effective teaching and learning practices are
2. provide effective modeling for teachers regarding effective teaching and learning practices are
3. use research on teaching and learning to guide strategic planning for accomplishment of school goals are
4. plan effective activities and experiences which facilitate teachers' beliefs in their abilities to provide effective teaching and learning activities to their students are
5. use data collected from teacher observations to inform school-wide efforts for improving teaching and learning are
6. regularly perform effective observations of teachers are
7. stay abreast of current best practices for facilitating effective teaching and learning are
8. communicate needs and goals necessary to enhance effective instructional effectiveness to faculty are
9. provide experiences that foster and facilitate high levels of teacher motivation towards teaching and learning are
10. protect instructional time so that effective teaching and learning can take place
11. facilitate an atmosphere that provides fair and consistent discipline for all students are
12. maintain healthy school/community relations are
13. maintain a school-wide atmosphere that is conducive to teaching and learning are
14. buffer teacher from unnecessary paperwork

THE EFFECT OF TEACHING TECHNOLOGY ON THE PERFORMANCE AND ATTITUDES OF ACCOUNTING PRINCIPLES STUDENTS

Homer L. Bates, University of North Florida
Bobby E. Waldrup, University of North Florida

ABSTRACT

Improvements in computer technology and audio-video equipment have allowed accounting faculty to significantly change the way accounting information is delivered to students. Presentation software, such as PowerPoint, allows faculty to build well-designed slides that would be impossible with the traditional chalk, blackboard and overhead method. The research question examined in this paper is whether PowerPoint presentations significantly increase student performance and attitudes in the principles of accounting course. The results display no significant statistical differences in either student performance or attitudes between those sections taught using PowerPoint and those taught using the traditional method.

INTRODUCTION

Computer technology allows an accounting faculty member to change the delivery method of the accounting information provided to students. Recently developed teaching technology enables faculty to include graphics, scanned images, animation, sound, and access to Internet web sites. Presentation software, such as PowerPoint, allows faculty members to build slides in a well-designed format that would be virtually impossible to duplicate using the chalk and blackboard method. However, this new teaching methodology does not come without considerable cost. There are enormous costs for hardware, software, and faculty training and preparation time. Considering the substantial direct and indirect costs of "teaching" technology, a very important question is "Do teaching technologies actually help students learn?" The major objectives of this research project were to determine whether a particular teaching methods results in measurable performance effects among accounting principles students and whether there is an effect in the attitude of the students toward the accounting profession and toward their instructor.

PREVIOUS RESEARCH

Recently the pressure on university administrators and faculty to increase the level of technology in the classroom has intensified. Referred to as an academic technological arms race (Jackson, 2000), universities are routinely expected to compete for rankings such as Yahoo! Internet Life's "most wired colleges." This pressure has become a centerpiece of debate on college campuses as can be seen by looking at virtually any edition of the *Chronicle of Higher Education* in the past five years (see for instance http://chronicle.com/infotech/wiredcampus.xml for historical references to these articles).

Previous research, primarily in the field of education, has demonstrated the potential of multimedia to enhance the learning process. Shank (1995) argued that multimedia-based instruction theoretically should promote student learning and retention. Many others (Pea and Gomez, 1992; Drook, 1994; Roberts, 1994; Liebowitz and Letsky, 1996; Jategaonkar and Babu, 1995) have developed successful applications of multimedia technology in the learning process.

Liu and Beamer (1997) provided anecdotal evidence that suggested positive pedagogical effects of multimedia in the classroom due to the ability to "attract, hold the attention, and spark the imagination of" the viewer. However, Becker and Watts (1998) suggest that most academic professors have not kept pace with this technology sufficient to employ these advantages. From evidence in the economics field, Greenlaw (1999) found that the only definitive outcome was an increase in the initial workload of faculty for new course preparations.

Over a decade ago, the Accounting Education Change Commission (1992, 1993) provided guidance that recommended, among other things, the use of materials in the first accounting course that "enhance presentation … consistent with current developments and new technology in the field…." Specific to the field of accounting education, Jensen and Sandlin (1992a and 1992b) strongly supported the use of technology in accounting education and provided information for those faculty interested in using multimedia approaches in the classroom.

Research into the effect of multimedia in the accounting classroom is very limited. Landry et al. (1997) reported that ten percent of the faculty at teaching universities uses multimedia and only four percent of those at research institutions use it. The most common type of multimedia used was videos. Rand et al. (1997) found that a multimedia presentation method had a significant effect on students' test scores immediately following a classroom presentation. This positive effect disappeared after two weeks. The multimedia presentation did not affect the students' evaluation of the topics or of the faculty members.

The *Chronicle of Higher Education* (Young, 2004) recently summarized the most comprehensive study to date regarding technology and teaching. In a survey conducted by the Educause Center for Applied Research, 4,374 students at 13 colleges of all types provided their perceptions of technology in the classroom. According to the report, 48.5 percent of the respondents said the biggest benefit of classroom technology is convenience, while only 12.7 percent of the

students said improved learning was the greatest benefit. Researchers who conducted the study specifically asked students to comment on professors' use of PowerPoint slide shows. Generally, the respondents were negative complaining that "faculty tend to read PowerPoint slides rather than teaching from them," leading the interviewer to conclude that PowerPoint used badly makes a lecture *worse*.

The only two business-specific studies which isolate the effects of computer-generated slide presentations (PowerPoint) upon student performance and interest were conducted by Hagen et al. (1997) in strategic management, and Rankin & Hoaas (2001) in the field of economics. Hagen et al. conducted an experiment with management students to determine if the inclusion of computer-aided presentations affected students' satisfaction, participation, and performance and found that all three variables were positively affected by the presentations. In a more specific experiment, Rankin and Hoaas (2001) observed 69 Principles of Economics students across two semesters and four class sections, half of which were exposed to PowerPoint presentations and half of which were exposed only to the traditional lecture method. Their results found no statistical effect on student performance, student attitudes towards the subject matter, or student evaluations of the instructor. Since the results of these two studies are contradictory and neither involved accounting courses, a well-controlled study focusing on accounting education would be beneficial.

HYPOTHESES

Drawing from this inconclusive literature, this experiment is designed to test hypotheses relating course presentation method to the performance and attitudes of accounting principles students. Thus, the model tested is:

Student Performance, Student Attitude f (Course Presentation Method) + e *Formula (1)*

Stated in the null, these hypotheses are:

H1: Student performance in an accounting principles course will not be affected by the instructor's presentation method of course material.

H2: Student attitude in an accounting principles course will be not be affected by the instructor's presentation method of course material.

Where:

Student performance is measured by individual examination results and grade distributions;
Student attitude is measured by course evaluations and additional survey results, and
Instructional presentation method is either primarily "chalk and talk" or Microsoft PowerPoint based delivery methods.

Since PowerPoint controls 93% of the presentation software market (Rankin & Hoaas, 2001), this study utilized it as the vehicle of study within the context of performance and attitudes in the first accounting course.

METHODOLOGY

An experiment was conducted in four accounting principles sections across two consecutive semesters. In this study, two types of teaching methods were compared: (1) traditional lecture using a blackboard for illustrations and transparencies on an overhead projector for problems and exercises, and (2) multimedia lecture using PowerPoint for lectures, problems and exercises. Prior to each of these lectures, copies of all slides were made available to the students in the PowerPoint sections only.

The experiment was conducted at a comprehensive regional (13,000 students) public university in the Southeastern United States. The sophomore level course, entitled "Principles of Financial Accounting," is the first accounting course offered, and is required of all Business majors as well as students majoring in Dietetics, Construction Management and Information Systems. The multimedia materials used in the course were developed during the summer term preceding the experiment, and the courses were taught during the following spring and fall academic terms. The PowerPoint presentation package accompanying the course text was modified and customized for the course.

The classroom in which the four sections were held was equipped with chalkboards, a traditional overhead machine, and "technology podium" at the front of the room. The podium contained a computer, VCR, DVD, and document camera. Multimedia presentations could be projected on a wide screen which could be lowered over the chalkboard.

During the first experimental academic term, the two course sections were taught in back-to-back time blocks in the same classroom. The first section was taught using the traditional blackboard approach, and the second section was taught using PowerPoint. During the second experimental term, the order was controlled by again teaching two sections consecutively, with the PowerPoint section taught first and the traditional section taught second. Each of the paired sections covered identical material with the same instructor, the same textbook and identical examinations. In addition, the classrooms for the paired sections were the same, the syllabi were identical, the quizzes were identical, and the grading scale was the same for both. The controls used in this study are shown in Table 1.

There were initially 154 students enrolled in the four test sections, of which 99 received final course grades, resulting in an attrition rate of 35.7%. This attrition rate is typical of the course, and did not significantly differ between the sections studied or from the departmental norm during, before, and after the study period.

The experiment was conducted during semesters when 20 Principles sections (11 in the fall and 9 in the spring) were taught by nine separate instructors. Aside from a common textbook, instructors are completely independent with respect to teaching and testing styles. The students in the experimental sections were not informed that the professor was utilizing a different teaching approach in the match-paired control sections.

Table 1: Controls Used in Study	
1	Same Instructor
2	Same Textbook
3	Same Syllabi and Assignments
4	Same Examinations
5	Same Classroom
6	Same Grading Scale

HYPOTHESES TESTING

Student Performance

Four examinations were given at equally spaced intervals throughout the test semesters. The three exams given during the term had a maximum score of 200 points each. The comprehensive final exam had a maximum score of 250 points. Each exam was comprised of approximately 50% multiple choice questions and 50% problem-based exercises. These four examinations totaled 85% of the final course grade.

Hypothesis one was first tested by comparing the mean scores for each exam between the match-paired course sections. T-tests were performed on the four examinations given during the fall term in the traditional section and the four examinations in the "PowerPoint" section to observe whether there was a significant difference in the means. This test was repeated during the spring term. The results of these eight t-tests are shown in Tables 2 and 3. As shown in the Tables, none of the t-statistics was significant at the .05 level of significance. There were no significant differences between the mean examination scores for any of the eight pairs of examinations given during the experimental fall and spring terms.

The final grade distributions in the two pairs of classes were also examined using the chi-square test. The grades given in each of the sections were in the traditional "A,B,C,D, & F" format. The chi-square statistic measures whether there is a difference between sections in the distribution of these grades. These results are shown in Table 4. There was no significant difference between the grade distributions in the two pairs of classes. The traditional approach and the PowerPoint

approach resulted in similar final grade distributions (The grade distributions for Principles of Accounting were examined for terms prior to the two-semester research period and subsequent to the research period. The distribution of grades during these periods was not significantly different from the grade distributions during the research period.).

Table 2: Student's t-Test Results of Exam Scores - Fall – Semester 1

	N=	Mean	Standard deviation	95% confidence interval for mean	t=	Probability=
Exam 1						
Traditional	44	131	21	124-138	1.00	0.32
PowerPoint	38	136	23	129-143	1.00	0.32
Exam 2						
Traditional	30	122	34	107-136	0.606	0.55
PowerPoint	31	128	46	114-142	0.606	0.55
Exam 3						
Traditional	22	141	27	128-155	0.774	0.46
PowerPoint	21	134	36	120-148	0.774	0.46
Final Exam						
Traditional	22	180	32	162-198	1.33	0.19
PowerPoint	21	163	51	144-182	1.33	0.19

Given that neither individual examination scores nor overall grade distributions were significantly different between the experimental and control class sections, we fail to reject hypothesis one. Accordingly, no evidence could be found that student performance in the course was differentially affected by the inclusion or exclusion of PowerPoint presentations.

Student Satisfaction

At the end of each semester, students in each section were given two forms of satisfaction surveys to complete. The first survey was developed by the instructor and consisted of six attitudinal questions about accounting as a profession / subject matter in general. The six questions were asked using a five-point Likert scale and are shown in Table 5.

Table 3: Student's t-Test Results of Exam Scores - Spring – Semester 2						
	N=	Mean	Standard deviation	95% confidence interval for mean	t=	Probability=
Exam 1						
Traditional	38	144	26	135-153	1.10	0.28
PowerPoint	37	137	31	127-146	1.10	0.28
Exam 2						
Traditional	30	129	27	118-139	0.026	0.979
PowerPoint	30	128	22	118-139	0.026	0.979
Exam 3						
Traditional	29	150	27	140-159	0.238	0.813
PowerPoint	28	151	22	142-161	0.238	0.813
Final Exam						
Traditional	27	178	32	164-192	0.934	0.354
PowerPoint	29	169	51	155-182	0.934	0.354

Table 4: Chi-square Test Results for Grade Distribution			
Courses compared using the Chi-square Test on A, B, C, D, F, W Distribution	Degrees of Freedom	Chi-square Statistic Computed	Significant at the 95% Confidence Level or Above
Fall Term PowerPoint/Traditional	5	2.81	No Critical Value: 11.07
Spring Term PowerPoint/Traditional	5	1.79	No Critical Value: 11.07

Table 5: Questions Asked to Determine Student Attitudes						
		Strongly Disagree				Strongly Agree
1	I hate accounting	1	2	3	4	5
2	I enjoy thinking about accounting.	1	2	3	4	5
3	I am no good at accounting.	1	2	3	4	5
4	Accounting is a useful subject.	1	2	3	4	5
5	Accounting has little appreciation in my life.	1	2	3	4	5
6	Accounting is an exciting subject.	1	2	3	4	5

Table 6 presents the Chi-square test utilized to determine between-section differences in these attitudes. There were no significant differences between the distribution of responses for the paired sections for either the fall or the spring experimental terms.

Finally, the effect of the delivery system (PowerPoint) on the students' evaluations of the faculty member was examined. Since 1996, the applicable State Legislature enacted legislation mandating that all State University System faculties administer the eight-question State University System Student Assessment of Instruction (*SUSSAI*) form in every section of every course taught (Chancellor's Memorandum, 1995). The results are public information. The mandatory SUSSAI questionnaire is shown as Table 7.

Table 6: Chi-square Test Results for Attitude Questions

Questions Compared	Degrees of Freedom	Chi-Square Statistic Computed	Significant at the 95% Confidence Level or Above
Fall – Semester 1			
Question 1	4	4.61	No*
Question 2	4	2.77	No*
Question 3	4	0.71	No*
Question 4	4	4.09	No*
Question 5	4	1.90	No*
Question 6	4	4.73	No*
Spring – Semester 2			
Question 1	4	6.06	No*
Question 2	4	5.04	No*
Question 3	4	1.66	No*
Question 4	4	4.87	No*
Question 5	4	2.77	No*
Question 6	4	2.66	No*
* Critical Value is 9.49			

Specific instructions are given regarding the administration of this common evaluation form. It must be administered at the beginning of class near the end of the term, with the faculty member being evaluated absent during the evaluation period. Results of the evaluations are not made available to the faculty member until after course grades have been assigned.

Table 8 shows chi-square test results for the paired sections for the fall and the spring terms. During the fall term, students viewed the PowerPoint section more positively than the traditional

section. The distribution of the student evaluations was significantly different for Questions 1, 2 and 8. For all three questions, the evaluations of the students in the PowerPoint section were higher. On Question 1 – Description of course objectives and assignments, over 44% of the students in the PowerPoint section responded "Excellent," while none of the students in the traditional section responded likewise. For Question 2 – Communication of ideas and information, over 22% of the students in the PowerPoint section responded "Excellent," while none of the students in the traditional section responded likewise. For Question 8 – Overall rating of instructor, over 27% of the students in the PowerPoint section responded "Excellent," while none of the students in the traditional section responded "Excellent."

For the spring term, there were no significant differences between the two sections. When fall and spring were combined, there were no significant differences between the distributions in the two types of sections. While there may have been a mild effect upon student evaluations of the instructor, it appears, however, that the delivery method did not significantly and consistently affect the student's evaluation of the instructor of the course.

Based upon the combined results of the two satisfaction instruments across all four sections of the course, we fail to reject the null of hypothesis two. Accordingly, little evidence could be found that student satisfaction in the course was differentially affected by the inclusion or exclusion of PowerPoint presentations.

Table 7: State University System Student Assessment of Instruction (SUSSAI)

Directions: Please assess your instructor's performance on the following eight items by darkening one response for each. Respond to each item according to the CODE printed below:

E = Excellent	VG = Very Good	G = Good	F = Fair	P = Poor	NR = No Responese

1	Description of course objectives and assignments.	E	VG	G	F	P	NR
2	Communication of ideas and information.	E	VG	G	F	P	NR
3	Expression of expectations for performance in this class.	E	VG	G	F	P	NR
4	Availability to assist students in or out of class.	E	V	G	F	P	NR
5	Respect and concern for students.	E	V	G	F	P	NR
6	Stimulation of interest in the course.	E	V	G	F	P	NR
7	Facilitation of learning.	E	V	G	F	P	NR
8	Overall rating of instructor.	E	VG	G	F	P	NR

CONCLUSION

The education literature is replete with anecdotal but conflicting evidence that technological innovations in the classroom affect student performance and attitudes about the subject matter being presented. While the debate will continue as to the direction and strength of these effects, the objective of this study is to lend empirical support to this discussion.

Specifically, the sophomore level Principles of Accounting course was examined to test whether the inclusion of PowerPoint presentations affects student performance and attitudes. In this study, no discernible effects could be supported that PowerPoint presentations have any impact on student outcomes.

While the study controlled for many possible factors, it did not control for the individual faculty member. Different faculty members may find differences because of their unique teaching style. It should also be noted that this study only examined the use of PowerPoint slides. It is possible that the use of other means of delivery and course management such as videos, Internet connections, Blackboard, WebCT, etc., could result in significant differences.

The results of this study do, however, call into question the rush to bring more and more technology into the classroom. As illustrated in the introduction, technology is very expensive, both in terms of outlay costs and opportunity costs. For instance, in the authors' university, traceable costs for the business building alone approach $500,000 per year.

In this study, existing PowerPoint slides were adapted for classroom use. An average of over ten hours per week was expended in adapting these existing slides. Preparing one's own slides would obviously require a significantly greater time commitment. Perhaps it is time to stand back and carefully examine the effect of these new, costly delivery systems. If they do not make a difference, then the question is "Why incur the additional cost?"

Table 8: CHI-SQUARE TEST RESULTS FOR SUSSAI QUESTIONS			
Questions Compared	Degrees of Freedom	Chi-Square Statistic Computed	Significant at the 95% Confidence Level or Above
Fall – Semester 1			
Question 1	4	11.40	Yes*
Question 2	4	11.80	Yes*
Question 3	4	6.13	No*
Question 4	4	2.83	No*
Question 5	4	6.41	No*
Question 6		.5	No*
Question 7	4	6.87	No*
Question 8	4	13.75	Yes*
Spring – Semester 2			
Question 1	4	3.62	No*
Question 2	4	3.56	No*
Question 3	4	2.93	No*
Question 4	4	2.36	No*
Question 5	4	0.36	No*
Question 6	4	4.60	No*
Question 7	4	2.07	No*
Question 8	4	1.03	No*
Fall & Spring Semester Combined			
Question 1	4	6.42	No*
Question 2	4	8.79	No*
Question 3	4	0.93	No*
Question 4	4	3.42	No*
Question 5	4	4.62	No*
Question 6	4	5.18	No*
Question 7	4	5.57	No*
Question 8	4	8.49	No*
* Critical Value is 9.49			

Table 9: Classroom Technology Costs	
Direct Cost	
Podium	$65,000
Hardware	$19,500
VCR/DVD	$5,200
Projector's A.V. piece	$110,500
Camera	$39,000
Bulbs	$16,250
Software licenses	$1,300
Total Direct Costs	$256,750
Allocated Costs	
T ech support	$250,000
Total Associated Costs	$506,750

REFERENCES

Accounting Education Change Commission (1992). The first course in accounting: Position Statement no. 2. *Issues in Accounting Education, 7*(2), 249-251.

Accounting Education Change Commission (1993). Evaluating and rewarding effective teaching: Issues Statement No. 5. *Issues in Accounting Education, 8*(2), 436-439.

Agarwal, R. & Day, A.E. (1998). The impact of the internet on economic education. *Journal of Economic Education,* (Spring), 99-110.

Becker, W. & Watts, M. (1998). Teaching economics: what was, is, and could be. *Teaching Economics To Undergraduates: Alternatives to Chalk and Talk*. Northampton, MA: Edward Elgar Press, 1-10.

Chancellor's memorandum 95-06.1, State of Florida State University System, June 1995.

Drook, G. (1994). Transcript of keynote address Washington interactive multimedia 1994, *Journal of Instructional Delivery Systems,* (Fall), 3-9.

Greenlaw, S. (1999). Using groupware to enhance teaching and learning in undergraduate economics. *Journal of Economic Education,* (Winter), 33-42.

Hagen, A., Edwards, K., & Brown, A. (1997). The impact of technology on classroom presentations: an empirical investigation. *The Journal of Private Enterprise,* (Spring), 54-68.

Jackson, G. (2000). The wired campus: enough is enough. *The Chronicle of Higher Education,* (August 11) B8-B9.

Jategaonkar, V.A., & Babu, A.J.G. (1995). Interactive multimedia instructional systems: a conceptual framework. *Journal of Instructional Delivery Systems,* (Fall), 24-29.

Jensen, R.E. & Sandlin, P.K. (1992a). Why do it? advantages and dangers of new waves of computer-aided teaching/instruction and what software features to consider. *Journal of Accounting Education,* (Spring), 25-38.

Jensen, R.E. & Sandlin, P.K. (1992b). How to do it? getting started in authoring computer-aided teaching/instruction and what software features to consider. *Journal of Accounting Education,* (Spring), 39-60.

Landry, R.M., Case, T., & Francisco, W. (1997). Multimedia in accounting: a survey of its use and relationship to a school's teaching/research mission. *1997 Proceedings of the Annual Meeting of the Southeastern American Accounting Association,* 181-191.

Liebowitz, J. & Letsky, C. (1996). KARMA: The knowledge acquisition reference multimedia aid. *T.H.E. Journal,* (February), 85-87.

Liu, D. & Beamer, L. (1997). Multimedia as a teaching tool in business communication course delivery. *Business Communication Quarterly,* (June), 51-66.

Pea, R.D. & Gomez, L.M. (1992). Distributed multimedia learning environments: why and how? *Interactive Learning Environments,* (February), 73-109.

Rand, R.S., Galbreath, S.C., & Snodgrass, P. (1997). Marriage or divorce: Can multimedia and accounting unite to enhance student performance? *1997 Annual Meeting of the Southeastern American Accounting Association*, 192.

Rankin, E. L. & Hoaas, D. (2001). Teaching note: does the use of computer-generated slide presentations in the classroom affect student performance and interest? *Eastern Economic Journal*, (Summer), 335.

Roberts, B.C. (1994). Information highways: delivery and shaping the multimedia world of tomorrow. *Executive Speeches,* (April/May), 4-6.

Shank, R. Active learning: integrating multimedia and artificial intelligence. *Keynote Address at the International Symposium on Artificial Intelligence*. Monterrey, Mexico, October 18, 1995.

Young, J. R. (2004). Students say technology has little impact on teaching. *The Chronicle of Higher Education,* (August 13), A28.

ACTIVE LEARNING: AN EMPIRICAL STUDY OF THE USE OF SIMULATION GAMES IN THE INTRODUCTORY FINANCIAL ACCOUNTING CLASS

Lou Fowler, Missouri Western State University

ABSTRACT

One of the major criticisms of accounting education today is that it is too oriented toward memorization and procedure. Employers and others complain that it produces graduates that are incapable of critical thinking and independent evaluation. The traditional lecture based delivery of material, followed by exercises in recording data, is thought to exacerbate this problem. This paper compares the use of the traditional lecture format using debit/credit journaling to the use of an active learning format using case analysis and a simulation game. The author's hypothesis is that the active learning format encourages higher level skills on Bloom's taxonomy which involve critical thinking. If this is true, it could help point the way is which accounting educators could address the needs of future graduates.

Both pedagogies are tested using a quiz constructed to follow Blooms taxonomy of learning. The quiz has questions that range from the ability to recall knowledge up to the ability to evaluate and make judgments about the validity of how material is gathered and how it is presented. Two introductory financial accounting classes were designated as either the control class, which received the traditional lecture based format, or the treatment group which received the active learning format. Both classes received the same quiz, had the same instructor and the same textbook. The study is repeated in the next semester. A total of 101 students were involved.

Contrary to what the evaluator was expecting, there were no significant differences between the two groups on either the lower or higher level scales of Blooms taxonomy. However, the traditional pedagogy was superior in the middle scale, which involves the application of rules and procedures.

One cannot help but wonder if the ability to think critically may be one of those skills that needs to be developed at an early age, and if it isn't, can become very difficult to learn later. Or could it be that students, who are proficient in memorization and prefer procedural methods, are more likely to choose accounting as a major? Answers to these questions, and more like them, are crucial in addressing the criticisms that are being leveled at accounting education today and preparing our graduates for successful entry into the business world.

INTRODUCTION

Practice sets, debits/credits and the techniques of the accounting system are still the primary focus of accounting educators in the introductory financial accounting class today. Although one would guess that these time honored techniques are effective. There has been little empirical evidence to support that. There has also been a growing concern that those techniques produce graduates that are good at memorizing and processing data but not effective in developing graduates with critical thinking and independent evaluation skills. Accounting educators, practitioners and students agree that accounting education has become increasingly procedure and knowledge oriented. These sentiments and others like them calling for change in accounting education have come from highly respected sources.

Albrecht and Sach (Albrecht and Sach, 2000) in their monograph prepared for the American Accounting Association have made dismal predictions about the course of accounting and accounting education in the future. According to extensive surveys and interviews that Albrecht and Sach conducted, "Accounting leaders and practicing accountants are telling us that accounting education, as it is currently structured, is outdated, broken and needs to be modified significantly." Current teaching pedagogy is specifically targeted as being problematic. Overemphasis on memorization and recall and too much reliance on lectures and textbooks is perceived as nothing more than a "trained monkey" approach. Creative types of learning such as team work, case analysis, role playing, writing assignments and out-of classroom experiences are not being used enough.

In July 2000, the AICPA, American Institute of Certified Public Accountants commissioned The Taylor Research and Consulting Group, Inc. to conduct an extensive research study among high school and college students to find out; 1. The decision-making processes of students in choosing college majors and career choices 2. General attitudes and perceptions of the accounting profession 3. Ways to encourage students to consider and choose the accounting profession (The Taylor Research & Consulting Group, Inc., 2000).

According to the report, the most pressing problem facing the accounting profession today is that it is not attracting enough students. The research indicates that the biggest hurdle seems to be with misperceptions about what accountants do. Most students' still perceive accounting to be largely a desk job requiring a lot of paper shuffling and math. The image of the thick glasses, green eyeshade and ten-key calculator still looms large. This is exactly the type of image that attracts students who prefer memorization and procedures and repels students who prefer creativity and critical thinking.

The Taylor report goes on to say that not only do accounting educators need to change their pedagogy but they also need to market those changes to potential students at an early age, preferably by high school. The marketing should also extend to academic advisors. Most advisors recommend that students take high schools bookkeeping as a preparatory course for college accounting. This

reinforces the concept that accounting is mostly a procedural oriented field. Advisors need to inform students that the accounting profession requires graduates with good communications skills, are independent problem solvers and have the ability to adapt to a constantly changing business environment.

According to the Accounting Education Change Commission (AECC, 1990), the curriculum for general education should develop students that have the capacities for inquiry, abstract logical thinking and critical analysis. The AECC further identified (AECC, 1992) that the first introductory financial accounting class should have as its' objective the requirement that students be able to critically analyze and solve unstructured problems. As a result accounting educators have begun to develop techniques that put the emphasis on student learning through active participation, as opposed to teaching to a passive audience. It is estimated that 94% of AACSB, American Association of Colleges and Business Schools, programs had active learning as a component of the curriculum (Faria, 1990). In general, active learning, or experiential learning as it is sometimes called, may be a way to rejuvenate accounting education.

RELATED RESEARCH AND HYPOTHESIS

In recent years accounting educators have made many efforts to respond to the demands for change. A great deal of that research has focused on different types of active learning techniques. Barkman (Barkman, 1998), Hassall (Hassall,1998) studied the effects of using case analysis to promote critical thinking. (Craig and Amernic, 1994) used role-playing in an effort to develop higher level thinking in students. Beets (Beets, 2003) studied the effects of learning in groups. Alder and Milne researched the motivation behind problem-based learning (Alder and Milne, 1997). McEwen (McEwen, 1994) evaluated case studies and problem solving as a more effective teaching method for developing critical thinking skills.

How to measure critical thinking has in itself been a topic of much discussion. The most influential and widely recognized work in critical thinking is the *Taxonomy of Educational Objectives* (Bloom, et al. 1974). Bloom's Taxonomy, as it is referred to today, was developed by a committee of college and university examiners and published as two handbooks- *Cognitive Domain* and *Affective Domain.* The categories listed in Bloom's Taxonomy range from what is referred to as lower level thinking skills such as knowledge and comprehension to the higher level skills such as synthesis and evaluation, Figure 1.

In an effort to develop and measure critical thinking in the classroom, most educators strive to ask questions that are structured around the higher scales of Blooms Taxonomy. Researchers have used many forms of active learning techniques to try to stimulate and measure those qualities (Burns, et al. 1990). It is presumed that passive learning, through reading and lectures are thought to produce lower level thinking, while interactive techniques like those found in active learning produce higher level thinking skills (Gentry, 1990).

Figure 1: Bloom's Taxonomy: Cognitive Domain	
Knowledge:	Knowledge of specific facts, terminology, conventions, classifications and categories.
Comprehension:	The ability to translate, interpret, and extrapolate.
Application:	The use of abstractions in the form of general ideas, rules or methods which must be remembered and applied.
Analysis:	Analysis of elements, relationships and organizational principles.
Synthesis:	Production of a unique communication, plan, or proposed set of operations. Derivation of a set of abstract relations.
Evaluation:	Judgments in terms of internal evidence and external criteria.

Jerry Gosenpud (Gosenpud, 1990) has done a great deal of work in reviewing the literature about active learning research and has published an evaluation that summarizes the results. Summary reviews were also done by Keys and Wolfe (Keys and Wolf, 1990). According to these reviewers, the studies taken as a whole were inconclusive. However that could be because few of these studies were designed to assess active learning effectiveness and many lacked well defined statistical controls.

To try and shed some light on the subject, the author designed a study to capture the quantitative differences between active and passive learning in developing critical thinking and evaluation skills. The null hypothesis is that active learning will yield the same results as traditional learning in increasing higher order thinking. The alternative hypothesis is that active learning yields better results in improving higher order thinking.

METHODOLOGY:

According to Albrecht (Albrecht, 1995), two of the most effective active learning techniques in developing critical thinking skills, are case studies and simulation games. This paper examines the difference between the uses of these two active learning techniques in the classroom to the traditional debit/credit accounting system.

At a Midwestern University, 101 students agreed to participate in a study that was conducted in the Introductory Financial Accounting class. The study spanned two semesters. In each semester one class was designated as the treatment group and received the active learning techniques and the other class as the control group which received the traditional lecture based method. There was a total of 52 students in the treatment group and 49 students in the control group. The goal for each group was to learn the difference between cash based accounting and full accrual accounting. All classes had the same teacher, the same textbook, and were in the same size and type of classroom. The average ACT in all four groups of students ranged from 20.5 to 21. The average grade in the

four groups ranged from 2.5 to 2.65. Each group received the same post-test. No pre-test was given because the topic tested over was new material for each group.

The treatment group received a case that involved a short scenario that covered a full accounting cycle (one month) of a lawn-mowing service. This was first done on a cash basis by using play money similar to that used in Monopoly. Groups of 4-5 students were formed and the group assumed the role of the business owner. The instructor took the position of the various other parties that the owner conducted business with and cash actually changed hands as the transactions occurred. The students were then asked to discuss the business' success within the group and come up with a consensus as to how well the business had fared.

The students were given the same case again. This time no cash was involved and instead the groups were asked to use full accrual accounting and use the accounting equation format. This required the students to analyze the transactions and identify which, if any, needed to be recorded. They had to determine what accounts would be involved, and how that would effect the accounting equation and the financial statements. The class then discussed the results of the full accrual method and how it differed from the cash basis method. To measure learning each student was asked to independently complete a ten point quiz (shown in Appendix A). The quiz was created by the instructor and contained questions that were constructed to follow each domain of Bloom's Taxonomy. There were two multiple choice questions each for Knowledge, Comprehension, Application, and Analysis and one open ended question each for synthesis and evaluation. Each question was worth one point.

Students in the control group were given the traditional lecture delivery of the same concepts, cash basis accounting and accrual basis accounting. As usual students were shown some exercises from the text and these were solved by recording debit/credit entries in the general journal format. Each student then completed the same quiz that had been given to the treatment group. Neither group was aware that the other group was involved until after the quizzes were returned in the next class period and the students were debriefed.

RESULTS

The conventional break between lower level thinking skills and upper level thinking skills is an even split. Knowledge, comprehension and application are considered lower level and analysis, synthesis and evaluation are considered upper level. As can be seen from Table 1, there was very little difference between the means for the active learning/treatment group and the traditional lecture/control group. Therefore, the null hypotheses, that active learning will yield the same results as traditional learning in increasing higher order thinking, cannot be rejected.

Table 1 Descriptive Statistics for performance on Bloom's Taxonomy						
	Lower Level Thinking Skills			Upper Level Thinking Skills		
	Mean	SD	N	Mean	SD	N
Treatment	.54	.30	52	.38	.14	52
Control	.53	.29	49	.35	.11	49
Differences in Means	.01			.03		

Further analysis among all of the domains does reveal a significant difference in the application domain, as can be seen in Table 2. However the stronger group appears to be the traditional group which received the lecture method and then solved problems using the more procedural oriented debit/credit journaling approach. According to Bloom, the application domain involves the use of abstractions in the form of general ideas, rules or methods which must be remembered and applied. This embodies the basic concepts in the debit/credit, double entry method of journalizing. So, it is not surprising that the traditional group would perform better in that area.

Table Two Descriptive Statistics for all Domains of Bloom's Taxonomy			
Domain	Active Learning N = 49	Traditional N = 52	Difference Between Means
	Mean	Mean	
Knowledge	.86	.78	.08
Comprehension	.52	.39	.13
Application	.25	.43	-.18
Analysis	.41	.39	.02
Synthesis	.34	.31	.03
Evaluation	.36	.33	.03

When performing t-tests on all six domains, refer to Table 3, it can be seen that the only statistically significant results were confirmed in the comprehension and application domain. Even though the knowledge domain did not pass the t-test for statistical significance, there does seem to be some degree of difference between the two groups. In all areas of Bloom's Taxonomy the active learning group outperformed the traditional group (although not to a significant degree in the upper level domains) with the exception of the application area. The negative results in the application domain indicate that the traditional group fared far better than the active learning group.

Table Three : T test results for all Domains of Bloom's Taxonomy						
t-test	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
T stat	1.52	1.89	-2.79	.25	.33	.35
T critical 1 tail	1.65	1.65	1.65	1.65	1.66	1.66
Significance	No	Yes	Yes	No	No	No

DISCUSSION AND CONCLUSIONS:

Although this author found no significant differences between active learning and traditional passive learning in the student's critical thinking and evaluation skills, there are some interesting results in the data on the lower level domains of Bloom's Taxonomy. It is confirmed that the active learning group scored higher in the comprehension domain and almost confirmed that they scored higher in the knowledge domain. These domains indicate the use of thinking skills that are largely accomplished through rote learning in which memorization is key. It has been shown that the more senses you use in memorizing, the better the results. The active learning techniques used in this study required that the students actually handle money, interact in a simulated business environment, discuss the concepts verbally and write down their conclusions thus utilizing more sensory input. In order to have scored well on the application domain, the traditional group also had to memorize the basic concepts. However, they were not able to learn these concepts as well as the active learning group. The traditional group did do significantly better in the application domain indicating that although they may not have as strong a grasp of the material, they certainly knew the accounting rules and procedures required to eventually put the material into appropriately coded format, such as the income statement, that is used in accounting. This indicates that they focused more on the application of the coding and formatting of the data, than actually comprehending what information was being conveyed. Unfortunately, this is exactly what today's critics of accounting education are referring to when they complain about today's accounting graduates being too procedure and knowledge oriented. However, one does need to keep in mind that knowing how to apply the accounting rules appropriately is a necessary part of accounting. Therefore, the traditional teaching method does have a place in today's educational environment, but shouldn't be the only method used. Perhaps, a blend of different teaching methods would yield better results in the lower level domains of Bloom's Taxonomy.

As to the question of which method improves higher level thinking, it appears that neither did in this study. To coin an old adage "if the answer was that obvious, we would have figured it out a long time ago". It wasn't until Bloom's Taxonomy was published in 1974, that we even had an acceptable construct to refer to, and even Bloom's has its critics. This study involved students that were in the first introduction to accounting class, which comprises mainly second semester

freshman or first semester sophomores. Perhaps critical thinking skills develop over time as younger people begin to confront complex issues more independently. To learn to think critically and be able to make independent evaluation, students need to be in an environment where there is no teacher or mentor readily supplying the answers. Internships, independent study courses and applied learning experiences, which are typically taken much later in the curriculum might be where one would expect to find evidence of higher level thinking.

One of the flaws of this study is that it involved students early in their curriculum and at only one point in their study. One cannot help but wonder if a longitudinal study covering the same groups of students over an extended period of time might yield different results. However, those types of studies are often difficult to arrange. The logistics involved in tracking the same groups over several semesters is daunting. Also, by its very nature active learning does not lend itself well to rigorous research design. Random selection of treatment and control groups, standardized pre-and post-tests, student-teacher relationships, class size and classroom features are often difficult to control and even more difficult to measure.

The majority of the studies reviewed by Gosenpud (cited earlier in this paper), found no difference between the active learning and the traditional pedagogies. However, one questions as to whether that is more a result of pitfalls of research design or is indeed a true statement. According to Gosenpud, "it should be tentatively concluded that in university classrooms the active learning method is no better or worse than other teaching methods for enhancing cognitive learning".

This study does highlight one striking point though, that the procedural knowledge based system of teaching accounting students how to input data into the system, does indeed work. Our traditional pedagogy in accounting education does help students to memorize the rules and procedures required to record business events in a particular manner. The problem of course is in the ability to critically analyze those events and project how those events, and changes in those events, will impact business outcomes. I also think that one would agree that the traditional approach stops short of addresses the needs of students graduating in such a complex world as today's and that more dynamic and varied learning approaches need to be explored.

REFERENCES

Accounting Education Change Commission (AECC). (1990). Objectives of education for accountants: Position statement number one. *Issues in Accounting Education* (Fall): 307-312.

Accounting Education Change Commission (AECC). (1992). The first course in accounting: Position Statement Number Two. *Issues in Accounting Education* (Fall): 249-251.

Albrecht, W.S., and R.J. Sack. (2000). *Accounting Education: Charting the Course through a Perilous Future.* Accounting Education Series (16). Sarasota, FL: AAA.

Alder, R.W., and M.J. Milne. (1997). Improving the quality of accounting students' learning through action-oriented learning tasks. *Accounting Education* 6(2): 191-215.

Beets, S. D. (2003). Differences In Student Leaning Techniques: Cooperative Learning, Quizzes, And Homework. *Journal of the Academy of Business Education* (Fall) 1-11.

Bloom, Benjamin S. and others. (1974). *The Taxonomy of Educational Objectives: Affective and Cognitive Domains*. New York: David McKay Company, Inc.

Burns A. C., J.W. Gentry, and J. Wolfe. (1990). A Cornucopia of Considerations in Evaluating the Effectiveness of Experiential Pedagogies. In Gentry, J.W. (Eds.), *Guide to Business Gaming and Experiential Learning.* (253-278). East Brunswick NJ: Nichols/GP Publishing.

Barkman, A. I. (1998). The use of live cases in the accounting information systems course. *Journal of Accounting Education* 16 (3): 507-516.

Craig, R., and J. Amernic. (1994). Role playing in a conflict resolution setting: Description and some implications for accounting. *Issues in Accounting Education* (Spring): 28-44.

Faria, A. (1990). Business simulation games after thirty years: Current usage levels in the united states. In Gentry, J.W. (Eds.), *Guide to Business Gaming and Experiential Learning.* (36-47). East Brunswick NJ: Nichols/GP Publishing.

Gentry, J.W. (1990). What is Experiential Learning? In Gentry, J.W. (Eds.), *Guide to Business Gaming and Experiential Learning.* (9-20). East Brunswick NJ: Nichols/GP Publishing.

Gosenpud, J. (1990). Evaluation of Experiential Learning. In Gentry, J.W. (Eds.), *Guide to Business Gaming and Experiential Learning.* (301-329). East Brunswick NJ: Nichols/GP Publishing.

Hassall, T., S. Lewis, and M. Broadbent. (1998). Teaching and learning using case studies: A teaching note. *Accounting Education* 7 (4): 325-334.

Keys, B. and Wolfe, J. (1990). The Role of Management Games and Simulations in Education and Research, *1990 Yearly Review of Management* (Summer), 16 (2), 307-336.

McEwen, B. C. (1994). Teaching Critical Thinking Skills in Business Education. *Journal of Education for Business.* 7 (2): 99-103.

The Taylor Research & Consulting Group, Inc. (July 2000). Student & Academic Research Study. *Presented to the AICPA*. Retrieved August 23, 2005, from http://ftp.aicpa.org/public/download/members/div/career/edu/Taylor_Report.pdf

APPENDIX A

10 point quiz

Aneta Break has decided to start a new computer repair business. She hires you as her accountant and a college student to work part-time. You have the following transactions for the month.

On January 2, Aneta writes a check on her personal bank account of $70,000 and asks you to deposit it in the bank in the company name of PC's R Us. On January 3, Anita buys equipment costing $15,000. She pays $5,000 down on the equipment with the promise to pay the remainder on February 3. On January 5, Aneta authorizes you to pay $5,000 for the first months rent.

During the month of January she and the student work on 500 computers at a cost of $40 each. One fourth of that is paid in cash and the rest is due by February 10. To encourage business she starts a service to clean and check hard drives for free. You prepare the payroll and pay the monthly wages of $20,000 to employees.

Use the above information to answer the following questions.

		Level on Bloom's Taxonomy
1.	Revenues are recorded when a. cash is received b. at the end of each month c. when work is performed d. at the discretion of the owner	Knowledge
2.	Expenses are a. the commitment to pay b. when cash is paid c. costs that occur in businesses d. costs of generating revenues	Knowledge
3.	What is the cash balance at the end January? a. $45,000 b. $60,000 c. $55,000 d. $35,000	Comprehension
4.	What is the amount of revenue earned during the month of January? a. $5,000 b. $20,000 c. $5,000 d. $15,000	Comprehension
5.	What amount of net income or net loss did PC's R us make? a. $20,000 income b. $5,000 income c. $5,000 loss d. $20,000 loss	Application

APPENDIX A		
6.	What would be the result if the company's accounts receivable were collected and the accounts payable paid. a. The company would increase income by $5,000. b. The company would increase assets by $5,000. c. The company would increase cash by $5,000. d. There would be no change in net assets.	Application
7.	Which of the following would cause income on the accrual basis to be different than income on a cash basis? a. Provide $5,000 services to a customer for cash. b. Borrowed $10,000 from the bank c. Owner invested $100,000 in the company. d. Charged $1,000 operating expenses on account.	Analysis
8.	Why do you think accrual basis accounting is preferable to cash basis accounting in evaluating a business's operating results? a. It isn't. Cash basis accounting is preferable, because you have to have money to conduct operations. b. It is preferable because it recognizes the economic benefits and the economic costs from operating activities when they occur and one can then compare the benefits of operations against the costs. c. It isn't. Cash basis accounting is preferable because it nets the cash received from operating activities against the cash paid for operating activities. d. It is preferable because it recognizes all of the revenues derived from a business and all of the expenses.	Analysis
9.	Referring back to the original scenario, how can Anita change her business activities to best improve income?	Synthesis
10.	Referring back to the original scenario, how would you evaluate Anita's first year of business operations?	Evaluation

Answers:
1.C 2.D 3.A 4.B 5.C 6.C 7.D 8.B 9 and 10. Answer varies

A CONFIRMATORY FACTOR ANALYSIS OF THE SCHOOL ADMINISTRATOR EFFICACY SCALE (SAES)

Daniel L. McCollum, University of Houston-Clear Lake
Lawrence T. Kajs, University of Houston-Clear Lake
Norma Minter, University of Houston-Clear Lake

ABSTRACT

Efficacy refers to peoples' beliefs and confidence to execute actions to attain a specific goal. Although knowledge of efficacy is well developed regarding students' learning and teachers' success, there is almost no research on efficacy of school administrators. However, McCollum, Kajs, and Minter (2006) developed a model and measure of school administrators' efficacy. The scale was based on the Educational Leadership Constituent Council (ELCC) national standards. Through exploratory factor analysis, eight dimensions of school administrator efficacy were derived. They included (1) Instructional Leadership and Staff Development, (2) School Climate Development, (3) Community Collaboration, (4) Data-based Decision Making Aligned with Legal and Ethical Principles, (5) Resource and Facility Management, (6) Use of Community Resources, (7) Communication in a Diverse Environment, and (8) Development of School Vision. This scale is titled School Administrator Efficacy Scale (SAES). In the present study, a model of the SAES was tested using a confirmatory factor analysis and hypothesizing the eight dimensions of the scale posited by its original researchers. The hypothesized model was tested using a sample of 559 principals and principal trainees. This research advances the knowledge and measurement of school administrators' efficacy, as additional evidence of the scale's validity and reliability is provided.

INTRODUCTION

Consider two assistant principals early in their careers, Carl and Sandra. Both have graduated with Master's degrees from respectable institutions. From their course work, one could assume that both have the necessary knowledge, skills, and dispositions to perform their new jobs. Same skill set, same level of experience, yet in short time, Sandra is outperforming her peer, Carl. Sandra works better with the community, is more effective in handling diversity, more adept in addressing instructional issues, and overall makes better decisions in the administration of her school. What sets these two assistant principals apart? Carl often finds himself thinking that he cannot get the job done, while Sandra stands tall and is confident in the knowledge, skills, and dispositions she has gained. In terms of Social–Cognitive Theory (Bandura, 1986), Sandra is self-efficacious, while Carl

lacks self-efficacy. Efficacy refers to "peoples' judgments of their capabilities to organize and execute courses of action required to attain designated types of performances" (Bandura, p. 391). Efficacy can impact one's performance range on a variety of tasks as well as one's motivational level for self-improvement to better engage in effective school practices (Schultz & Schultz, 1998). These elements are especially important for carrying out responsibilities of instructional leadership, since "Principal quality is linked statistically and practically to student achievement" (Kaplan, Owings, & Nunnery, 2005, p. 43).

The present article briefly summarizes self-efficacy as studied in students and teachers. From there, the pertinent, though scarce, literature on school administrators' (e.g., principals') efficacy is reviewed. Then, a recently developed model of a school administrator efficacy scale is posited. The model is theoretically based on Bandura's (1986) Social-Cognitive Theory, and its associated measure stems from the Educational Leadership Constituent Council's (ELCC's) national standards. The theoretical model has been empirically tested by McCollum, Kajs, and Minter (2006) and is further tested, along with its measure – the School Administrator Efficacy Scale (SAES) – in the present article.

SELF AND TEACHER EFFICACY

Research supports the claim that self-efficacy is a key construct in education. Self-efficacy is related to persistence, effort, and success on tasks (Bandura, 1986). Covington (1984) asserts that one's successes can lead to greater efforts to accomplish and persevere through difficult assignments. Thereby, achievement on tasks brings about self-efficacy, which in turn leads to greater success. In addition, self-efficacy can be increased through vicarious experiences (Bandura). That is, seeing the successful actions of others in their field can influence persons' beliefs that they too can master similar tasks (Bandura). Therefore, observing their mentors' effective performances can work to build up protégés' self-efficacy.

The importance of the efficacy construct has been made apparent as it relates to students and increasingly as it applies to teachers (e.g., Tschannen-Moran, Woolfolk-Hoy & Hoy, 1998). Henson, Kogan, and Vacha-Haase (2001) have concluded that several positive behaviors of teachers (e.g., ability to deal more effectively with failing students, persistence when difficulties arise) are linked to teacher efficacy, which positively impact student outcomes. Self-efficacy research, however, has mostly surrounded teachers and students, with little knowledge of the construct's use to understand the actions of school administrators, e.g., principals. Given the benefits of being efficacious (e.g., greater effort, persistence, and success), this construct may have extensive promise in the study of administrators who work in schools.

SCHOOL ADMINISTRATOR EFFICACY

More study is needed about the implications of efficacy on school administrators. For instance, do efficacious principals more effectively handle job stress, as well as relationships with staff, students, and parents? Do school administrators, who demonstrate high efficacy, employ more effective management practices, e.g., organizational planning, problem solving, and community building? These issues are critical in the face of the changing roles of school administrators and recent changes in their preparation.

Compared to the traditional approaches to school administrator preparation, five major shifts seem to have occurred for the current preparation of school administrators, including an emphasis on (a) interpersonal skills, (b) consensus development, (c) accountability processes, (d) integration of community and school needs and resources, and (e) policy development (NPBEA, 2002-a). It is understandable that communication has become a major focus since principals can spend up to 80 percent of their time on communicating with students, campus staff, parents, and the larger community (Green, 2001). The necessity in having school administrators who can demonstrate effective communication and social skills to address conflict resolution and consensus building situations in the school and larger community is supported by field research on principals (Kajs, Decman, Cox, Willman, & Alaniz, 2002). Because of the added attention to state and national accountability mandates, e.g., No Child Left Behind (NCLB), the shift to have principals more thoroughly prepared in accountability processes is also understandable.

Research is needed to determine efficacy levels in potential school administrators as well as in current ones. Two studies on principal efficacy include Dimmock and Hattie (1996) who found efficacy as a valued element for principals in a school restructuring process, and Smith, Guarino, Strom and Adams (2006**)** who concluded that the quality of teaching and learning is influenced by principal efficacy. Some principal efficacy measures that have been developed include the (1) Principal Self-Efficacy Survey (PSES) (Smith et al.); (2) Principals Sense of Efficacy Scale (PSES) (Tschannen-Moran & Gareis, 2004); (3) principal efficacy vignettes (Dimmock & Hattie); and (4) School Administrator Efficacy Scale (SAES) (McCollum, Kajs, & Minter, 2006). Smith et al. measured efficacy in Instructional Leadership and efficacy in Management. Those authors offered construct validity evidence in the form of factor analysis. Though their measure may be promising in terms of validity, it only captures two dimensions of the principal's job. More thorough in their investigation were Tschannen-Moran and Gareis (2004) who tested multiple measures of principals' efficacy. Those authors concluded that Dimmock and Hattie's (1996) measure was neither valid nor reliable and, therefore, could not be used in further studies or in practice. However, Tschannen-Moran and Gareis found some factorial validity and reliability for the scale they created – the Principal Sense of Efficacy Scale (PSES). Those authors captured three dimensions of the principal's job (i.e., management, instructional leadership, and moral leadership). Though their instrument is promising in terms of its psychometric properties and it expands beyond the work of

Smith et al., there may be potential to further capture the administrator efficacy construct by identifying additional dimensions of the job. McCollum, Kajs, and Minter developed a scale to measure school administrators' (e.g., principals) efficacy. Those authors noted the construct (factorial) validity of the scale and high reliability coefficients for its subscales. Through exploratory factor analysis, eight dimensions of school administrator efficacy were derived. These dimensions included (1) Instructional Leadership and Staff Development, (2) School Climate Development, (3) Community Collaboration, (4) Data-based Decision Making Aligned with Legal and Ethical Principles, (5) Resource and Facility Management, (6) Use of Community Resources, (7) Communication in a Diverse Environment, and (8) Development of School Vision (McCollum, Kajs, & Minter).

The scale created by McCollum, Kajs, and Minter (2006) was based on the ELCC national standards. These standards incorporate the well-known Interstate School Leaders Licensure Consortium (ISLLC) standards (Murphy, 2005). ELCC's leadership framework provides a roadmap for university-based educational administrator preparation programs regarding specific knowledge, skills, and depositions related to key themes in the development of school principals and superintendents (NPBEA, 2002-a). The work of Kaplan, Owings, and Nunnery (2005) noted the link between principal quality and ISLLC (ELCC) standards. Results indicated that principals who demonstrated higher ratings on these standards were leaders of schools with higher achievement among students; in contrast to school administrators who scored lower (Kaplan, Owings, & Nunnery, 2005). Their study noted that competent teachers want to work with effective principals, not ineffective ones; thus, the quality of leadership can directly impact teacher retention levels (Kaplan, Owings, & Nunnery).

The current ELCC Standards consists of seven standards toward the preparation of school administrators. Standards one through six (1-6) address the chief components of school administration (e.g., community communications and collaboration), while standard seven (7) focuses on applying and synthesizing content, skills, and dispositions outlined in standards one through six (1-6) through an internship experience. The seven ELCC Standards are outlined in Table 1 (NPBEA, 2002-b, pp. 1-18):

The SAES and its eight dimensions provide the model and measure for the present study. McCollum, Kajs, and Minter (2006) cite that this scale can serve as a useful tool in the development of future and current school leaders since subscales address knowledge, skills, and dispositions incorporated in the ELCC Standards, especially since there are a few studies related to principals' efficacy and its measurement.

Table 1. Seven ELCC Standards	
Standard 1	Candidates who complete the program are educational leaders who have the knowledge and ability to promote the success of all students by facilitating the development, articulation, implementation, and stewardship of a school or district vision of learning supported by the school community.
Standard 2	Candidates who complete the program are educational leaders who have the knowledge and ability to promote the success of all students by promoting a positive school culture, providing an effective instructional program, applying best practice to student learning, and designing comprehensive professional growth plans for staff.
Standard 3	Candidates who complete the program are educational leaders who have the knowledge and ability to promote the success of all students by managing the organization, operations, and resources in a way that promotes a safe, efficient, and effective learning environment.
Standard 4	Candidates who complete the program are educational leaders who have the knowledge and ability to promote the success of all students by collaborating with families and other community members, responding to diverse community interests and needs, and mobilizing community resources.
Standard 5	Candidates who complete the program are educational leaders who have the knowledge and ability to promote the success of all students by acting with integrity, fairly, and in an ethical manner.
Standard 6	Candidates who complete the program are educational leaders who have the knowledge and ability to promote the success of all students by understanding, responding to, and influencing the larger political, social, economic, legal, and cultural context.
Standard 7	The internship provides significant opportunities for candidates to synthesize and apply the knowledge and practice and develop the skills identified in Standards 1-6 through substantial, sustained, standards-based work in real settings, planned and guided cooperatively by the institution and school district personnel for graduate credit.

THE NEED FOR VALIDITY EVIDENCE

Given that efficacy is an understudied construct in the domain of school administration, a need exists to develop high quality instruments to measure the construct, as well as to study the construct further using such instrumentation. Establishing the validity of a measurement instrument is a key process in the development of good instrumentation. Benson (1998) offers three stages of construct validation: (1) substantive, (2) structural, and (3) external. In the substantive stage, constructs are theorized and defined. In the structural stage, relationships among variables purported to measure the construct are sought. Such techniques as exploratory and confirmatory factor analysis and internal consistency measures (e.g., Cronbach's Alpha) are utilized. The external stage incorporates the construct's relation to other constructs (i.e., creating the nomological network, see Cronbach & Meehl, 1955). This study focuses on advancing the structural stage of Benson's plan for developing the construct validity of the SAES. Previous research (i.e., McCollum, Kajs, &

Minter, 2006) has addressed the substantive stage and has only begun to address the structural stage of construct validation.

PURPOSE

Previous research has established initial evidence of the reliability and construct validity of the SAES (i.e., McCollum, Kajs, & Minter, 2006). Still, further evidence is needed to support the scale's construct validity; in previous work only exploratory factor analytic techniques were used. This study tests the SAES model using a confirmatory factor analysis and hypothesizing the eight dimensions of the scale posited by its original researchers. This study is being conducted to determine the construct validity of the SAES, and to lead to improvements in the measurement of school administrators' efficacy. In addition, the internal consistency of the scale will be re-evaluated in this new sample, using Cronbach's Alpha. This study serves to advance the knowledge and measurement of school administrator efficacy.

METHOD

Participants

The study participants were early career principals and principal trainees (n = 559). The participants were teaching in school districts or carrying out principal functions in the Houston, Texas area. The participants' mean teaching experience was 7.8 years (SD = 5.72). The mean experience as a principal was 7.8 months (SD = .41). The participants' mean age was 34.8 years (SD = 7.77). The sample consisted of 79.4 percent females and 20.6 percent males. The study sample was 51.2 percent European-American, 22.9 percent African-American, 22.5 percent Hispanic, 1.3 percent Asian, 1.1 percent biracial people, .4 percent Latino, .2 percent Native American, and .4 percent other.

Instrumentation

The SAES uses 51 items (see Appendix for items) and is purported to measure eight dimensions of school administrators' efficacy, using a seven-point Likert-type scale (1 = not at all true of me, 7 = completely true of me). The eight dimensions and their reliability coefficients (Cronbach's Alpha) based on McCollum, Kajs, and Minter (2006) are (1) Instructional Leadership and Staff Development (.93), (2) School Climate Development (.93), (3) Community Collaboration (.91), (4) Data-based Decision Making Aligned with Legal and Ethical Principles (.93), (5) Resource and Facility Management (.89), (6) Use of Community Resources (.95), (7) Communication in a Diverse Environment (.81), and (8) Development of School Vision (.86). These dimensions were

derived through exploratory factor analysis. Hence, some initial evidence of construct validity exists. The content validity of the SAES instrument comes from its base in the Educational Leadership Constituent Council (ELCC) national standards. This scale is young, but given its initial validity and reliability evidence, there is promise that the research conducted in this current study will lead to clear construct validity evidence; thereby, furthering the measurement of school administrator efficacy.

PROCEDURE

The SAES was given in group administrations to the 559 principals and principal trainees in the sample. Participants filled out a consent form, acknowledging their participation, and were provided a set of instructions for completing the SAES and a copy of the instrument. The SAES took approximately 20 minutes to finish.

RESULTS

The hypothesized model for the school administrator efficacy dimensions (see Appendix for listing of items by dimension) was tested using confirmatory factor analysis in EQS 6.1. The normalized estimate of multivariate kurtosis was found to be high (180.71), so robust maximum likelihood estimation was used in parameter estimation. According to Hu, Bentler, and Kano (1992) in cases of high multivariate kurtosis, typical of item data, a robust estimation method is desirable for corrective purposes. Upon completion of parameter estimation, the Bentler-Bonnett nonnormed fit index (NNFI), robust comparative fit index (RCFI – a robust calculation of CFI), root mean squared residual (SRMR) and root mean squared error of approximation (RMSEA) were selected among fit indices to evaluate the model's goodness of fit. These indices performed well in simulation studies and yielded complimentary information (i.e., Hu & Bentler, 1998, 1999; Hutchinson & Olmos, 1998; Marsh, Balla & Hau, 1996). Furthermore, it is a composite of these four criterions that is typically used when evaluating model fit/testing a hypothesis using confirmatory factor analysis.

The NNFI of .90 and RCFI of .90 were at the .90 standard for acceptable fit given by Bentler (1992). The SRMR of .06 was in a good range for fit, below the acceptable point for model fit of .08 (seeking values less than .08). As well, the RMSEA of .05 was in a good range for fit, given the standard of less than or equal to .08. Hu and Bentler (1998, 1999) gave the standard for SRMR, and Browne and Cudeck (1993) gave the standard for RMSEA. All four criteria – NNFI, RCFI, SRMR and RMSEA – suggest that the model fits. Therefore, the evidence in support of the model is strong. The Path Coefficients, Error Variances, and Variances Accounted For (R^2), item Means and item Standard Deviations are presented in Table 2.

Table 2
Path Coefficients, Error Variances, R-squared, Means, and Standard Deviations

Item	Path Coefficients	Error Variance	R-squared	M	SD
1	.87	.49	.76	6.06	.95
2	.89	.46	.79	6.08	.89
3	.82	.57	.68	5.99	.99
4	.68	.73	.47	6.31	.83
5	.70	.72	.49	6.23	.82
6	.82	.57	.67	6.04	.99
7	.76	.63	.60	6.10	.91
8	.76	.65	.58	5.96	.97
9	.88	.48	.77	6.01	.03
10	.88	.48	.77	5.99	.97
11	.86	.50	.75	6.09	.94
12	.70	.72	.48	5.77	1.07
13	.70	.72	.49	5.99	1.02
14	.75	.66	.56	5.64	1.06
15	.75	.67	.56	6.07	2.67
16	.83	.56	.69	5.86	1.05
17	.81	.58	.66	6.16	.97
18	.81	.59	.65	6.09	.93
19	.68	.73	.46	5.86	1.23
20	.66	.75	.43	6.30	.83
21	.83	.55	.70	6.01	.95
22	.84	.54	.70	6.09	.91
23	.80	.59	.65	5.97	.98
24	.82	.57	.67	6.28	.85
25	.79	.61	.63	6.05	.99
26	.94	.35	.88	5.72	1.09
27	.98	.21	.96	5.92	2.70
28	.59	.81	.34	6.67	.63
29	.79	.61	.63	5.74	1.22
30	.82	.57	.67	6.15	1.00
31	.87	.49	.76	6.08	1.03

Table 2
Path Coefficients, Error Variances, R-squared, Means, and Standard Deviations

Item	Path Coefficients	Error Variance	R-squared	M	SD
32	.80	.60	.65	6.15	.96
33	.80	.64	.59	5.94	.97
34	.86	.58	.67	6.14	.89
35	.83	.56	.68	5.77	1.14
36	.78	.62	.61	5.57	1.17
37	.82	.67	.68	5.61	1.27
38	.85	.53	.72	5.87	1.07
39	.75	.66	.56	6.16	.97
40	.80	.60	.64	6.00	1.00
41	.76	.66	.57	5.86	1.03
42	.84	.55	.70	5.72	1.18
43	.83	.56	.69	5.56	1.21
44	.71	.71	50	6.13	.93
45	.76	.65	.58	6.29	.86
46	.66	.75	.43	6.32	.92
47	.81	.59	.66	5.64	1.22
48	.62	.78	.39	6.66	.59
49	.60	.80	.36	5.74	1.18
50	.88	.47	.76	5.79	1.09
51	.86	.51	.74	5.77	1.07

SPSS 12.0 was used to calculate reliability coefficients (Cronbach's Alpha) for the SAES subscales and the correlations among the subscales. All of the correlations are statistically significant at the $p = .01$ level; however, they are low enough to warrant the conclusion that the subscales are separate. The subscale means, standard deviations, correlations, and reliabilities are shown in Table 3.

Table 3
Means, Standard Deviations, Correlations, (and Reliabilities) of Subscales

Subscale	M	SD	1	2	3	4	5	6	7	8
1. ILSD	5.98	.80	(.94)							
2. SCD	6.06	.80	.74	(.93)						
3. CC	5.73	.95	.69	.62	(.85)					
4. DBDM	5.92	.86	.72	.65	.73	(.92)				
5. RFM	6.01	.88	.68	.61	.71	.78	(.90)			
6. UCR	6.20	.82	.55	.48	.69	.56	.56	(.86)		
7. CDE	6.23	.74	.74	.67	.73	.68	.67	.66	(.81)	
8. DSV	6.11	.80	.69	.74	.57	.56	.51	.48	.66	(.90)

Note. ILSD = Instructional leadership and staff development, SCD = School climate development, CC = Community collaboration, DBDM = Data-based decision making aligned with legal and ethical principles, RFM = Resource and facility management, UCR = Use of community resources, CDE = Communication in a diverse environment, DSV = Development of school vision

CONCLUSIONS AND IMPLICATIONS

The null hypothesis of non-model fit is appropriately rejected based on the criterion typically used with confirmatory factor analysis. Therefore, the research hypothesis that the eight-factor model of the SAES fits is accepted. Thus, there exists strong evidence of the construct validity of the SAES. Additionally, the correlations among the subscales are low enough to warrant a conclusion that discriminant validity exists (e.g., the subscales are separate). The reliability coefficients of the eight subscales of the SAES were found to range from good (.81) to excellent (.94). These are important steps in the successful measurement of school administrators' efficacy. The findings regarding the construct validity and reliability of the SAES are consistent with and a significant addition to past findings (e.g., McCollum, Kajs, & Minter, 2006). With these new findings, researchers and practitioners can confidently rely on the validity and reliability of the SAES in their work.

Since school administrators as instructional leaders are responsible for student achievement, they need to participate in meaningful assessment practices to address their professional growth needs (Kaplan, Owings, & Nunnery, 2005). The SAES, which incorporates ELCC knowledge, skills and dispositions, has multiple assessment applications in the preparation of school administrators and the ongoing professional development of school principals. First, university principal preparation programs can use the SAES as a formative and summative instrument (i.e., pretest, mid-test, and posttest) to measure the growth and development of candidates as they progress during the

program. Particularly, during the internship stage, the principal-mentor and candidate could use this scale as a diagnostic instrument to confirm areas of strength and to develop an action plan to address content, skill, and disposition needs.

Secondly, this scale can be used to evaluate the success of a principal preparation program, when coursework is aligned to the ELCC standards. For instance, the SAES is currently being used as one of a variety of strategies to monitor and evaluate the development of candidates in the University of Houston-Clear Lake's Collaborative Bilingual Administrator Training (CBAT) program and to gauge the program's curriculum and delivery mechanisms in competently preparing future principals. CBAT is a federally funded five-year project to prepare highly qualified bilingual school principals who can serve the needs of English Language Learners (ELLs) in the Houston Metropolitan area.

Thirdly, the SAES can be used as a self-assessment instrument for practicing school administrators, enabling them to review and reflect upon their own strengths and needs as efficacious professionals. Results from this reflective, self-monitoring, standards-driven procedure can provide the basis for an individualized improvement plan. The process can be especially useful since principals are usually expected to outline a yearly professional development program, underscoring the lifelong learning mind-set of professionals. The practice of self-assessment can reduce reluctance among educators to be evaluated because it places them in charge of the information about their individual needs or perceived weaknesses; thus, eliminating public exposure of shortcomings (Jackson, 2005). Individual management of personal information encourages educators to be candid in assessing their knowledge and skill levels, as well as dispositions. Moreover, this confidential practice, as well as the opportunity to choose the relevant training programs to address needs could serve to increase their motivational level to participate and achieve success in their educational experiences (Schultz & Schultz, 1998). The cognitive processes of self-evaluating, self-supervising, and self-motivating, along with goal-development, planning, attention management, implementation of learning approaches, and solicitation of assistance from others when necessary comprise self-regulated learning, which is a key element in becoming an effective learner and leader (Ormrod, 2003).

The research in the present article can be used in the development of school leaders, as the scale addresses knowledge, skills, and dispositions described in the ELCC Standards, applicable to the effective training of school administrators. The SAES instrument can also serve as a viable tool for school administrators to self-evaluate their own strengths and needs as professionals, providing direction toward their professional improvement. Consequently, the SAES has potential to be used for multiple purposes of evaluation. The SAES is especially pertinent, considering that few self-efficacy assessments designed specifically for school administrators are available (Lashway, 2003). Future research addressing the relationships to and comparisons with Tschannon-Moran's and Gareis's (2004) PSES should be made, because both the SAES and PSES are psychometrically strong. It should be determined if those two instruments provide the same or complimentary

information. Future research on the SAES should also be conducted to establish how the constructs it measures relate to other constructs, such as motivation and work performance. This would further the validity program described by Benson (1998). Overall, the SAES is promising as a tool to validly and reliably measure school administrator efficacy.

REFERENCES

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Benson, J. (1998). Developing a strong program of construct validation: A test anxiety example. *Educational Measurement: Issues and Practice, 17*(1), 10-17.

Bentler, P. M. (1992). On the fit of models to covariances. *Psychological Bulletin, 88*, 588-606.

Browne, M. W., & Cudeck, R. (1993) Alternative ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing Structural Equation Model* (pp. 132-162). Beverley Hills, CA: Sage.

Covington, M. V. (1984). The self-worth theory of achievement motivation: Findings and implications. *The Elementary School Journal, 85*(1), 1-20.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281-302.

Dimmock, C., & Hattie, J. (1996). School principals' self-efficacy and its measurement in the context of restructuring, *School Effectiveness and School Improvement*, *7*(1), 62-75.

Green, R. L. (2001). *Practicing the art of leadership: A problem-based approach to implementing the ISLLC standards.* Columbus, OH: Merrill Prentice Hall.

Henson, R. K., Kogan, L. R., & Vacha-Haase, T. (2001). A reliability generalization study of the teacher efficacy scale and related instruments. *Educational and Psychological Measurement, 61*(3), 404-420.

Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin, 112*, 351-362.

Hu, L., & Bentler, P. M. (1998). Fit indexes in covariance structure modeling: Sensitivity to underparametrized model misspecification. *Psychological Methods, 3*(4), 424-53.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 11-55.

Hutchinson, S. R. & Olmos, A. (1998). Behavior of descriptive fit indexes in confirmatory factor analysis using ordered categorical data. *Structural Equation Modeling, 5*(4), 344-364.

Jackson, L. (2003). *Teacher training: Is your staff development program working.* Retrieved February 17, 2006, from Education World Web site: www.education-world.com/a_admin/ teacher_training/teacher_training005.shtml

Kajs, L. T., Decman, J. M., Cox, C., Willman, E., & Alaniz, R. (2002). Addressing mentor-principals' lifelong learning needs in the preparation of principal candidates. In G. Perreault & F. C. Lunenburg (Eds.), The changing world of school administration (pp. 325-346). Lanham, MD: Scarecrow Press.

Kaplan, L. S., Owings, W. A., & Nunnery, J. (2005). Principal quality: A Virginia study connecting interstate school leaders licensure consortium standards with student achievement. *NASSP Bulletin, 89*(643), 28-44.

Lashway, L. (2003). *Improving principal evaluation.* Eugene, OR: Clearinghouse on Educational Management. (ERIC Document Reproduction Service No. ED 482347)

Marsh, H. W., Balla, J. R., & Hau, K. (1996). An evaluation of incremental fit indexes: A clarification of mathematical and empirical properties. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 315-353). Mahwah, NJ: Lawrence Erlbaum Associates.

McCollum, D. L., Kajs, L. T., & Minter, N. (2006). School Administrator's Efficacy: A model and measure. *Education Leadership Review, 7*(1), 42-48.

Murphy, J. (2005). Unpacking the foundations of ISLLC standards and addressing concerns in the academic community. *Educational Administration Quarterly, 41*(1), 154-191.

NPBEA (National Policy Board for Educational Administration). (2002-a). *Educational leadership constituent council (ELCC): Instructions to implement standards for advanced programs in educational leadership.* Fairfax, VA: National Policy Board for Educational Administration. Retrieved February 17, 2006, from Instructions to Standards for preparing ELCC Program Report at http://www.npbea.org/ELCC/

NPBEA (National Policy Board for Educational Administration). (2002-b). *Educational leadership constituent council (ELCC): New ELCC Standards – NCATE approved.* Fairfax, VA: National Policy Board for Educational Administration. Retrieved February 17, 2006 at http://www.npbea.org/ELCC/

Ormrod, J. E. (2003). *Educational psychology: Developing learners.* Upper Saddle River, NJ: Prentice-Hall.

Schultz, D., & Schultz, S. E. (1998). *Psychology & work today: An introduction to industrial and organization psychology (7th ed.).* Upper Saddle River, NJ: Prentice Hall.

Smith, W., Guarino, A., Strom, P., & Adams, O. (2006). Effective teaching and learning environments and principal self-efficacy. *Journal of Research for Educational Leaders, 3*(2), 4-23.

Tschannen-Moran, M., Woolfolk-Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *68*(2), 202-248.

Tschannen-Moran, M., & Gareis, C. (2004). Principals' sense of efficacy: Assessing a promising construct. *Journal of Educational Administration, 42,* 573-585.

APPENDIX	
Factor 1:	**Instructional Leadership and Staff Development.**
12.	I am confident in my understanding of the total instruction program in my school.
13.	I am able to understand the process of curriculum design, implementation and evaluation.
14.	I am confident in my knowledge of best-practice research related to instructional practices.
15.	I am able to develop a systematic process for mentoring teachers on my campus.
16.	I am confident that I understand and can communicate to staff the complex instructional and motivational issues that are presented by a diverse student population.
17.	I am confident in my skills to lead staff to understand and respect the diversity of our student population.
18.	I am confident that I can lead staff to appreciate the kinds of knowledge and skills students and their families can add to the learning process.
19.	I understand the development of a professional growth plan.
20.	I have a clear sense of my own personal development needs and the resources I can access to address those needs.
21.	I am confident in my skills to assess the staff development needs of a school.
22.	I am confident that I possess the skills needed to implement the effective use of resources so that priority is given to supporting student learning.
23.	I am confident in my skills to engage staff in the development of effective campus improvement plans that result in improved learning.
Factor 2:	**School Climate Development**
5.	I have the ability to engage students in the assessment of our school climate.
6.	I have the ability to assess school climate using multiple methods.
7.	I have the ability to engage staff in the assessment of our school climate.
8.	I have the ability to engage parents in the assessment of our school climate.
9.	I am confident that I know how to use data about our school climate to improve the school culture in ways that promote staff and student morale.
10.	I am confident that I know how to use data about our school climate to encourage appropriate student behavior.
11.	I am confident that I know how to use data about our school climate to support a positive learning environment.
Factor 3:	**Community Collaboration**
35.	I understand community relations models that are needed to create partnerships with business community, and institutions of higher education.
36.	I am confident in my ability to use marketing strategies and processes to create partnerships with business, community, and institutions of higher education.
37.	I can identify and describe the services of community agencies that provide resources for the families of children in my school.
38.	I am confident in my skills to involve families and community stakeholders in the decision-making process at our school.
49.	I am confident I can resolve issues related to budgeting.
50.	I am able to supplement school resources by attaining resources from the community.
51.	I am confident I can solicit community resources to resolve school issues.

APPENDIX	
Factor 4:	**Data-based Decision Making Aligned with Legal and Ethical Principles**
39.	I can make sound decisions and am able to explain them based on professional, ethical and legal principles.
40.	I am confident in my ability to understand and evaluate education research that is related to programs and issues in my school.
41.	I am confident in my ability to apply appropriate research methods in the school context.
42.	I can explain to staff and parents the decision-making process of my school district.
43.	I can explain to staff and parents how the governance process of my school is related to state and national institutions and politics.
44.	I am confident in my ability to examine student performance data to extract the information necessary for campus improvement planning.
46.	I can make decisions within the boundaries of ethical and legal principles.
47.	I am able to explain the role of law and politics in shaping the school community.
Factor 5:	**Resource and Facility Management**
29.	I am confident in my knowledge of legal principles that promote educational equity.
30.	I am able to provide safe facilities (building, playground) according to legal principles.
31.	In accordance with legal principles, I am confident I can find information to address problems with facilities.
32.	I am able to find the appropriate personnel to resolve facility-related problems.
33.	I am confident in my ability to identify additional resources to assist all of the individuals in my school.
Factor 6:	**Use of Community Resources**
25.	I am confident I could use community resources to support student achievement.
26.	I am confident I could use community resources to solve school problems.
27.	I am confident I could use community resources to achieve school goals
Factor 7:	**Communication in a Diverse Environment**
24.	I am confident in my skills to interact positively with the different groups that make up my school community.
34.	I am confident in my ability to lead my staff in involving families in the education of their children.
45.	I am confident in my communication abilities to lead in a variety of educational settings.
Factor 8:	**Development of School Vision**
1.	I am confident that I possess the skills to lead a school community in the development of a clear vision.
2.	I can develop a vision that will help ensure the success of all students.
3.	I am able to use strategic planning processes to develop the vision of the school.
4.	I am confident that I can establish two-way communication with stakeholders (staff, parents, students, community) in order to obtain the commitment necessary for implementing the vision for our school.

THE ROLE OF GENDER IN TEACHING EFFECTIVENESS RATINGS OF FACULTY

Jonathan Kohn, Shippensburg University
Louise Hatfield, Shippensburg University

ABSTRACT

This paper examines the role of gender and its influence on student ratings of faculty teaching effectiveness. The study recorded professor effectiveness ratings by 930 undergraduate students consisting of 472 females and 458 males. The results reveal several gender differences. Generally, female students rated faculty effectiveness higher than male students. When gender of faculty was considered, female students rated male faculty higher than male students, but did not rate female faculty higher than male students. Gender differences were also examined using an integrated model for student rating behavior. This model included the theories of motivation, grade leniency/stringency, and construct validity, which have been integrated into a structural equation model. Previous research generally treated these theories independently. The effect of gender on the role each of the competing theories was studied. Rating behavior was generally consistent between male and female students. Females seemed to exhibit lower academic expectations than males. The results of this study, in conjunction with previous studies, continue to show that significant bias exists in student ratings of faculty teaching effectiveness regardless of gender of students or faculty.

LITERATURE REVIEW

Many studies have been conducted to study the various factors that influence student ratings of professor effectiveness. Those that have focused on gender differences have revealed inconsistencies related to faculty evaluations. Some studies have shown higher ratings for instructors by females, though in some instances same sex preferences were found also (Ferber & Huber, 1975, Tieman & Rankin-Ullock, 1985). Others studies have shown little or no gender interactions (Elmore & LaPoint 1974, 1975, Wilson & Doyle 1976). Hancock, Shannon & Trentham (1993) considered gender and college disciplines (human sciences, liberal arts, etc) and found no uniform patterns. They did find female students rated instructors higher than did males. Feldman (1993) did an extensive analysis and found minor differences and inconsistent results, though students rated same sex instructors somewhat higher. Fernandez ,1997 reviewed the literature and concluded that gender differences were minimal with regard to rating of faculty. His own study

supported these conclusions, stating "…that the effect of student and faculty gender on teaching quality assessment is slight to non-existent".

Other factors also impact teaching evaluations. Academic success expectancy has been studied and found that there were small gender differences with females expectancy slightly less than males (Gigliotti & Seacrest, 1988). The effects of motivation on professor ratings are probably the most agreed upon systematic influence in student ratings of faculty. It has been demonstrated that student motivation, represented by student interest and course type (elective/required), plays a significant role in student ratings of professor effectiveness (Howard & Maxwell, 1980; Hoyt 1973; Marsh, 1984; Marsh & Duncan, 1992). Howard and Maxwell (1980, 1982) modeled the relationship between student motivation, student learning, expected grades, and student satisfaction with the instructor and field of study. These and other studies show that motivation and learning are more highly correlated with ratings of professor effectiveness than is expected grade with professor effectiveness. The authors conclude that student motivation drives the correlation between grades and student satisfaction with the instructor. Therefore, the correlation between grades and ratings of professor effectiveness is an expected artifact, rather than an indication of a direct relationship between grades and ratings of professor effectiveness. Using path analysis, Marsh (1984) also concluded that prior subject interest had a stronger impact on student ratings of various professor effectiveness characteristics than did grades. Additionally, simple classifications (required versus elective) and expanded categories of course type have been found to be significantly correlated with ratings of professor effectiveness (Aleamoni, 1981; Centra, 1993; Feldman, 1978; Marsh & Dunkin, 1992).

Construct validity theory proposes that student ratings reflect student learning and, therefore, measure professor teaching effectiveness. That is, higher student ratings for the instructor indicate greater student learning. Some studies have demonstrated that classes with the highest student ratings have also performed best on standardized final examinations in multi-section classes (Marsh & Roche, 1997). In addition to the earlier studies that provided the foundation for validity theory, numerous factor analytic studies have been conducted to investigate the validity of student ratings (Cashin 1988; Feldman 1989; Howard, Conway, & Maxwell 1985; Marsh 1984; Marsh & Duncan 1992).

Using SEM, Greenwald and Gilmore (1997a&b) found support for grade leniency theory by suggesting that only grade leniency allows for a negative workload à grade relationship. This relationship is explained by students' willingness to work harder in order to avoid very low grades. This negative relationship between workload and grades has been observed in other studies (Marsh, 1980). However, other explanations have been offered for the negative relationship, such as subject difficulty and student capability (McKeachie, 1997

In an effort to integrate the competing theories, SEM analysis was conducted by Hatfield and Kohn, 2004, which confirmed the presence of all the competing theories and their interactions.

HYPOTHESES

Based on the literature and the findings of several studies, we propose that there are differences in the way male and female students rate male and female faculty.

H1: Female students rate faculty higher than male students.

H2: Female students rate male faculty higher than male students

H3: Female students rate female faculty higher than male students

Theories of student rating behavior suggest a variety of hypotheses which form the basis of an integrated approach to study gender differences. Several theories have been offered in explanation of the positive relationship between grades and ratings of faculty (Greenwald & Gilmore, 1997a). Grade leniency suggests that grades directly affect ratings of faculty. Construct validity and motivation assert that a third variable (learning positively affects both grades and ratings, thus, resulting in a positive relationship between grades and ratings. While these theories have been studied extensively in the past, the effect of gender has not been studied when all three theories are present simultaneously.

Grade Leniency/Stringency

This theory suggests praise induces liking for the individual giving the praise (Aronson & Linder, 1965; Hatfield & Kohn, 2003). In the context of student ratings, praise is interpreted to be high grades and liking is translated into high faculty ratings. Grade leniency theory suggests that there is a causal relationship between expected grades and ratings of faculty. Further, Greenwald and Gilmore (1997a) suggest that there is a negative relationship (grade stringency) between students working hard and expected grades. In courses that have strict-grading policies students have to work hard in order to avoid very low grades, yet overall grades are still lower than in classes with easy-grading professors. These premises suggest the following hypotheses:

H4: The higher the expected grade, the higher the professor effectiveness rating.

H5: The higher the student effort (worked harder), the lower the expected grade.

Construct Validity

This theory suggests that high instructional quality induces high student learning, which results in higher grades and higher professor ratings (Cashin & Downey, 1992; Cohen, 1981; Feldman, 1976 & 1989; Marsh, 1984). Therefore, the following hypotheses are provided to evaluate construct validity:

H6: The higher the student learning, the higher the professor effectiveness rating.

H7: The higher the student learning, the higher the expected grade.

H8: The higher the professor effectiveness rating, the higher the student learning.

H9: The higher the worked hard rating, the higher the student learning.

Student Motivation.

This theory suggests that student motivation positively affects both grades and ratings of faculty, through student learning, thereby resulting in a positive correlation between grades and ratings of faculty (Aleamoni, 1981; Braskamp & Ory, 1994; Centra, 1993; Kohn & Hatfield, 2001; Marsh, 1984; Marsh & Dunkin, 1992). Student motivation results in more student learning and appreciation for the course and instructor, which leads to higher grades and higher professor effectiveness ratings. Researchers have identified two measures of student motivation: course-specific and general (Howard & Maxwell, 1980; Marsh, 1984). These indicators of student motivation will be examined in this study—student interest in the subject matter of the rated course and course type (major or elective, versus required or core course). Student interest is a course-specific measure, whereas course type is a general measure. The following hypotheses are designed to test the impact of student motivation in student rating behavior:

H10: The higher the student interest, the higher the student learning.

H11: Lack of choice in course (required or core courses) results in lower student learning.

H12: The higher the student learning, the higher the expected grade.

H13: The higher the student learning, the higher the professor effectiveness rating.

RESEARCH METHODS

The student rating survey contained eight items, which students rated on a six-point Likert scale: (1) strongly agree, (2) agree, (3), slightly agree, (4) slightly disagree, (5) disagree, (6) strongly disagree. The first six items were designed to examine professor effectiveness, with the sixth item being a global item. Student learning was assessed by item 7 and course specific student interest by item 8.

1. *The course requirements, including grading system, were explained at the beginning of the semester.*
2. *The professor provides feedback on exams and assignments.*
3. *The professor is willing to answer questions and assist students upon request.*
4. *The professor uses examples and practical applications in class, which aid in my understanding of the material.*
5. *The professor encourages students to analyze, interpret, and apply concepts.*
6. *The professor was effective teaching this course.*
7. *I learned a significant amount in this course*
8. *I am interested in the subject matter of this course.*

These items were selected based on past research, which suggests the desirability of global items that address professor effectiveness (#6) and student learning (#7) factors, and the need to control for student interest (#8) (Cashin, 1995). Items one through five address commonly used dimensions of professor effectiveness in student rating research (Braskamp & Ory, 1994; Cashin, 1995; Centra, 1993; Feldman, 1989; Marsh, 1991).

Students completed a student data sheet that contained demographic items, two grade-related items, and one general student motivation item: (1) The grade I expect to achieve in this course, (2) I worked harder in this course than in most of my other courses, and (3) course type. All response options were designed so that students could use opscan sheets to report their ratings. The scale for expected grade was: 1.A, 2.B, 3.C, 4.D, 5.F. The agree-disagree Likert scale noted above was also used for the 'worked harder' item. Five categories of course type were provided: A. required by major/minor, B. elective in major/minor, C. general education requirement, D. free elective, E. program core course. These items reflect commonly used measures in testing for grade leniency and motivation effects on student ratings of faculty (Greenwald & Gilmore, 1997a&b; Howard & Maxwell, 1980; Marsh, 1984).

SAMPLE AND PROCEDURES

Data were collected from students and professors in the three colleges (business, arts and science, and education) at Shippensburg University at the end of the first semester of the 1997-1998 academic year. Classes were included in the sample from professors volunteering and by request (in order to insure adequate representation from all colleges and departments), a mix of student classes (such as freshman and senior), and a mix of professor characteristics (such as gender, race, degree, and rank).

Nine hundred and thirty students, (472 females and 458 males) and 44 professors (17 females and 27 males) were included in the sample, with the largest percentage (51) of faculty in Arts and Sciences, and equal percentages in Business and Education. The largest percentage of students were seniors, 36 percent, followed by sophomores at 19 percent, juniors at 18 percent, freshmen at 14 percent, and graduate at 13 percent.

VARIABLES AND MEASURES

The professor effectiveness dependent variable is a composite measure, developed by averaging the ratings of the six professor effectiveness items. The reliability coefficient, alpha, for the composite professor effectiveness measure is 0.84. *Expected Grade* is both a dependent and independent variable, and is used directly as reported. *Student Learning, Student Interest,* and *Worked Hard* are also used as directly reported in the survey instrument. The *Course Type* categories were collapsed into a two-category independent variable: (1) major/minor/elective, and (2) required/core course. There are two measures of student motivation: *Course Type* and "*Student Interest* in the subject matter of this course". There are two measures of grade leniency: *Expected Grade* and "*Worked harder* in this course than in most other courses". *Student Learning*, a self-reported rating, is a construct validity measure.

The scales for five variables (professor effectiveness, expected grade, student learning, student interest, worked hard) were reversed so that interpreting the findings would be more consistent with the way these variables are typically referred to, e.g., low to high. For example, the higher the student learning rating, the more the student learned, etc. The course type variable is categorical, and, thus did not need to be reversed.

ANALYSIS AND RESULTS

Descriptive statistics (means, standard deviations, and correlations) for all the variables used in this study are provided in Tables 1, 2, and 3. The hypotheses will be tested on the within-class data using structural equation modeling (SEM) and the Amos 5.0 modeling software. While there are many goodness-of-fit statistics in SEM, this study will report three of the most popular measures

(CFI, NFI, Chi-square/df), with Comparative Fit Index (CFI) being the primary fit-statistic used in this study (see End Notes). Path coefficients are tested for significance using Critical Ratios (CR). Amos 5.0 reports both the CR's and the P values for each path so that level significance can be determined. In addition, comparisons of professor effectiveness will be made among the combined male and female, male only, and female only samples for professor effectiveness rating behavior. Average scores will be tested to see if there are differences among the groups using a difference of means test. Finally, a comparison will be made to determine if male and female students rate male and female faculty members differently. A one-way analysis of variance will be performed to see if there are any significant differences in rating behaviors.

Table 1: Descriptive Statistics Males and Females: Correlations, Means and Standard Deviations (N =930)

		Prof. Effect	Student Learn.	Student Interest	Expect. Grade	Worked Harder	Course Type	Mean	Standard Deviation
Prof. Effect.	Correlation	1	.565	.381	.350	.065	-.182	5.34	.653
	Significance		.000	.000	.000	.047	.000		
Student Learn.	Correlation		1	.570	.314	.196	-.141	5.00	.959
	Significance			.000	.000	.000	.000		
Student Interest	Correlation			1	.360	.053	-.174	4.75	1.224
	Significance				.000	.107	.000		
Expect. Grade	Correlation				1	-.111	-.079	4.17	.775
	Significance					.001	.016		
Worked Harder	Correlation					1	-.178	4.07	1.332
	Significance						.000		
Course Type	Correlation						1	1.26	.441

A comparison was performed to contrast average professor effectiveness rating scores among male and females combined, females only, and males only. (see Tables 4 & 5). A test for the difference of means was performed (assuming unequal variances) and a highly significant difference ($Z = 3.453$, $P < .000$) was found between Females only and Males only. Thus, female average professor effectiveness rating scores are significantly higher than males. These results support H1 and similar findings (Benz and Blau 1995) which also found female ratings higher than males.

Table 2: Descriptive Statistics Females Only: Correlations, Means and Standard Deviations (N= 472)

		Prof. Effect.	Student Learn.	Student Interest	Expect. Grade	Worked Harder	Course Type	Mean	Standard Deviation
Prof. Effect.	Correlation	1	.619	.433	.367	.053	-.191	5.41	.633
	Significance		.000	.000	.000	.255	.000		
Student Learn.	Correlation		1	.602	.332	.140	-1751	5.07	.970
	Significance			.000	.000	.002	.000		
Student Interest	Correlation			1	.417	.056	-.203	4.71	1.247
	Significance				.000	.223	.000		
Expect. Grade	Correlation				1	-.209	-.041	4.17	.775
	Significance					.000	.372		
Worked Harder	Correlation					1	-.206	4.11	1.307
	Significance						.000		
Course Type	Correlation						1	1.26	.441

Table 3: Descriptive Statistics Males Only: Correlations, Means and Standard Deviations (N=458)

		Prof. Effect.	Student Learn.	Student Interest	Expect. Grade	Worked Harder	Course Type	Mean	Standard Deviation
Prof. Effect.	Correlation	1	.506	.340	.317	.072	-.175	5.27	.665
	Significance		.000	.000	.000	.126	.000		
Student Learn.	Correlation		1	.543	.284	.251	-.107	4.94	.944
	Significance			.000	.000	.000	.022		
Student Interest	Correlation			1	.312	.051	-.144	4.78	1.200
	Significance				.000	.273	.002		
Expect. Grade	Correlation				1	-.023	-.119	4.08	.770
	Significance					.618	.011		
Worked Harder	Correlation					1	-.150	4.03	1.358
	Significance						.001		
Course Type	Correlation						1	1.26	.441

Table 4: Means and Variances for Faculty Effective Ratings for Females and Males combined, Females only, Males only			
Gender	N	Mean	Variance
Males and Females	930	5.340	.426
Females Only	472	5.412	.401
Males Only	458	5.265	.442

Table 5: Differences in Mean Faculty Effectiveness Ratings for Females only and Males only		
Differences	Z Score	P value
Females - Males	3.453	.000

An analysis was conducted to determine if the gender of either students or faculty played a role in rating of faculty effectiveness. Because of fewer missing data for this analysis, the total sample size was 936 instead of 930 students. The gender of the faculty was noted and a one way analysis of variance was conducted among the 4 combinations of male and female students and male and female faculty. Table 6 presents the averages and standard deviations of faculty effectiveness ratings, along with sample size. Table 7 presents the results of the analysis of variance indicating significant differences among groups ($P < .001$). Sheffe multiple comparisons were made between male and female student ratings for male and female faculty. Female students rated male faculty significantly higher than did male students (mean difference = .177, $P < .05$), supporting H2. There was no significant difference in female and male students ratings of female faculty (mean difference = .103). Therefore, H3 is not supported.

Table 6: Measures for Faculty Effectiveness Rating Scores by Gender Groups				
	Female Fac. Female Stu.	Female Fac. Male Stu.	Male Fac. Female Stu.	Male Fac. Male Stu.
Average	5.43	5.33	5.40	5.223
Standard Deviation	0.627	0.640	0.638	0.6767
Sample size	189	159	283	305

Table 7: ANOVA Table: Faculty Effectiveness Rating Scores by Gender groups					
	S.S.	D.F.	MSQ	F	Sig.
Between Groups	6.65	3	2.218	5.269	.001
Within Groups	392.30	932	.421		
Total	398.95	935			

NEW PERSPECTIVES ON STUDENT RATINGS

One of the problems with testing each of the above theories in isolation of each other is that intervening and moderating effects on the predicted relationships are not taken into account. Such effects may suppress or reinforce the predicted relationships. Thus, to accurately assess the presence of the theorized relationships, all the variables of interest need to be included in the same model. This section will integrate the findings predicted by the various theories, using structural equation modeling. The impact of student gender will then be studied to determine the role gender plays in rating behavior.

Following similar methodology of Hatfield and Kohn 2004, all of the predicted direct relationships proposed in the grade leniency/stringency, construct validity, and motivation theories were used to construct integrated structural models for males and females, males only, and females only. The initial analyses of these models are presented in left hand column for figures 1, 2, and 3. Analysis of the model reveals that the fit for all three models was very good (CFI = .94 for all 3 models) and resulted in R^2 s of .35 (M & F), .41 (F), and .28 (M) for professor effectiveness. However, several path coefficients were not significant. For each model, removal of these paths was evaluated by iteratively removing the path with the highest P value greater than .05, rerunning the model with the path deleted, and then inspecting the P values of the remaining paths for those P values that were greater than .05. The procedure stopped when all remaining paths were significant. The final models are presented on the right hand side of figures 1, 2, and 3.

As a result of these procedures, both Course Type à Student Learning (Motivation H12) and Professor Effectiveness à Student Learning (Construct Validity H8) were deleted from all three models (M & F, M, F). In addition, Worked Harder à Student learning (Construct Validity H6) was deleted from the female model. Inspection of the modification indices of the final models of both males and females, males only, and females only indicated that no additional paths would strengthen the model. These results are presented in the right hand column of Figure 1. All paths of the final models are significant, the fit is very good (CFI = 1.00 (F) and .999 (M & F, M)), and R^2 s are .35 (M & F), .41 (F) and .29 (M). While the R^2 s of the final models remained at their original levels, the fit indices of both models improved significantly. The R^2 for all models are quite strong. R^2 for females only is 46% higher than for males only. All three of these R^2 values are much higher than

usually reported in many studies. In addition, there is a high degree of consistency in the structural nature of all models with only Worked Harder à Student Learning linkage being omitted from the female only model.

Figure 1 Integrated Models for Males and Females

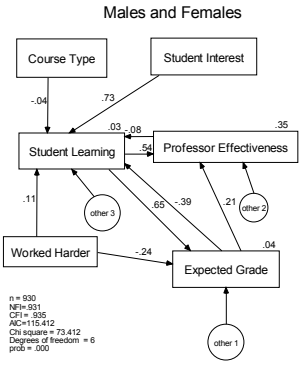


Figure 2 Integrated Models for Males

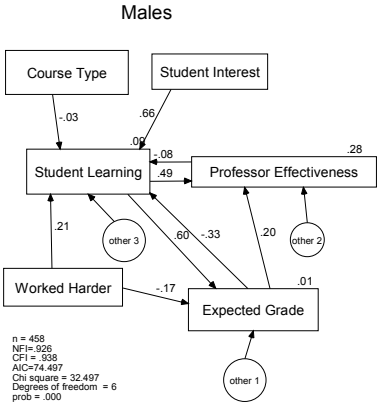
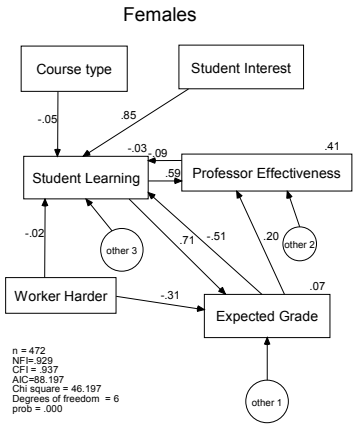


Figure 3 Integrated Models for Females



Thus, 9 of the original 12 hypotheses were strongly supported, lending considerable support to the three theories of student rating behavior, regardless of gender. Interestingly, both grade leniency (H4) and grade stringency (H5) are supported. While grade leniency is commonly understood and generally accepted, grade stringency (negative workload, expected grade relationship) has been rarely observed (Gilmore and Greenwald – 1997a & b). In this analysis, it is not only observed in the combined sample of males and females but in both sub-groups of male only and females only. Moreover, the standardized path coefficient between Worked Harder and Expected Grade is much stronger for females (-.31) than males (-.17).

In all final models, the Professor Effectiveness à Student Learning (H8) was removed. This hypothesis is part of Construct Validity theory and may indicate that the feedback loop between student learning and ratings of professor effectiveness may not be as well defined as assumed. A possible explanation for this weakness is that higher effectiveness ratings may not be a good measure of teaching ability and thus does not lead to greater student learning.

Although found in other studies, the Course Type (H12) link was dropped from all final models. It is generally assumed that students are more motivated in electives or courses in their major and less so in required courses. This facet of motivational theory then leads to higher professor effectiveness ratings. Our results do not indicate this to be so. Course Type is an indirect affect, influencing student learning which in turn affects professor effectiveness ratings. In an integrated model, Student Interest has a major impact (path coefficient - .70 (M&F), .81(F) and .64(M)), eliminating the role of Course Type component of Motivation theory. Thus, in an integrated model Course Type may be a redundant variable with Student Interest providing a much stronger indication of student motivation.

CONCLUSIONS

Some inconsistencies continue to show up in the study of gender on faculty effectiveness ratings. In general, females students rate faculty higher than did males students. When gender of faculty was considered, females students rated male faculty higher than did male students. Higher female rating scores have been observed in other studies and the results of our study strongly support this difference. However, the differences ended there. In addition, our findings also provide strong support for the need to integrate theories that explain student rating behavior of faculty. Integration is necessary because of the interactions and indirect effects among the theoretical premises. Structural equation modeling provides an ideal analytical methodology to study the complexities of student-rating behavior

Using an integrated approach based on SEM analysis, we have found surprisingly consistent results among all the models. Structurally, little difference in overall student rating behavior based on gender differences was observed. Except for the removal of one path for females (Worked Harder à Student Learning, H_6, Construct Validity), the three models are identical. All exhibit the identical simultaneous effects of Construct validity, Grade leniency and Stringency, and Motivational theories. All models fit the data very well and have much higher coefficients of determinations than previously reported. These high values continue to lend support for the significant bias that exists in rating of professor effectiveness that continues to be ignored by many schools. Females experience the negative workload expected grade effect to a much higher degree than males. Some studies have found that females have lower expectancy of success (Gigliotti and Seacrest 1988). The results of our study tend to support these findings. Faced with a difficult course, females may be less likely to have the confidence in themselves and thus expect a lower grade though they will continue to work harder.

This study continues to find significant bias, including gender bias, in student ratings of faculty, suggesting the need to re-consider how student ratings are used to evaluate faculty teaching effectiveness. In order to more accurately evaluate professor effectiveness, administrators and faculty need to control for, or at least acknowledge, the complexity student rating behavior.

ENDNOTES

A 1.0 CFI or NFI suggests a perfect fit and if under .9 the model can probably be improved (Bentler and Bonnett, 1980). Chi-square/df ratios of up to 3 are indicative of acceptable fit models (Marsh and Hocevar, 1985). CFI is less affected by sample size than is NFI or the Chi-square ratio (Kline, 1998).

REFERENCES

Aleamoni, L.M. 1981. Student ratings of instruction. In J. Millman (Ed.), *Handbook of teacher evaluation* (pp. 110-145). Beverly Hills, CA: Sage.

Anderson, J.C. & Gerbing, D.W. 1988. Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3): 411-423.

Aronson, E. & Linder, D.E. 1965. Gain and loss of esteem as determinants of interpersonal attractiveness. *Journal of Experimental Social Psychology*, 1: 156-171.

Bentler, P.M. & Bonnett, D.G. 1980. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88: 588-606.

Bentler, P.M. & Chou, C. 1987. Practical issues in structural modeling. *Sociological Methods and Research*, 16: 78-117.

Benz, C. & Blatt, S.J. 1995. Factors Underlying Effective College Teaching: What students Tell Us. *Mid-Western Educational Researcher*, 8 (1): 27-31.

Bollen, K.A. 1989. *Structural Equations with Latent Variables*. New York: Wiley.

Braskamp, L.A. & Ory, J.C. 1994. *Assessing faculty work: Enhancing individual and institutional performance*. San Francisco: Jossey-Bass.

Bridgeman, W.J. 1986. Student evaluations viewed as a group process factor. *Journal of Psychology*, 120: 183-190.

Cashin, W.E. 1988. Student ratings of teaching: A summary of the research. *Idea Paper No. 20*. Manhattan: Kansas State University, Center for Faculty Evaluation and Development.

Cashin, W.E. 1995. Student Ratings of Teaching: The Research Revisited *IDEA Paper No.32*. Manhattan: Kansas State University, Center for Faculty Evaluation and Development.

Cashin, W.E. & Downey, R. 1992. Using Global Student Rating Items for Summative Evaluation. *Journal of Educational Psychology, 84*: 563-572.

Centra, J.A. 1993. *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness.* San Francisco: Jossey-Bass.

Chacko, T.I. 1983. Student ratings of instruction: A function of grading standards. *Education Research Quarterly*, 8(2): 14-25.

Chapman, J.W. & Lawes, M.M. 1984. Consistency of causal attributions for expected and actual examination outcome: A study of the expectancy confirmation and egotism models. *British Journal of Educational Psychology*, 54: 177-188.

Cohen, P.A. 1981. Student ratings of instruction and student achievement. A meta-analysis of multi-section validity studies. *Review of Educational Research*, 51: 281-309.

D'Apollonia, S. & Abrami, P.C. 1997. Navigating student ratings of instruction. *American Psychologist*, 52(11): 1198-1208.

Davis, M.H. & Stephan, W.G. 1980. Attributions for exam performance. *Journal of Applied Social Psychology*, 10: 235-248.

Elmore, P.B. & LaPointe, K.A. 1975. Effects of teacher sex, student sex, and teacher warmth on the evaluation of college instructors. *Journal of Educational Psychology*, 67,368-374.

Feldman, K.A. 1976. Grades and college students' evaluations of their courses and teachers. *Research in Higher Education*, 4: 69-111.

Feldman, K.A. 1978. Course characteristics and variability among college students in rating their teachers and courses: A review and analysis. *Research in Higher Education*, 9: 199-242.

Feldman, K.A. 1989. The association between student ratings of specific instructional dimensions and student achievement: Refining an extending the synthesis of data from multi-section validity studies. *Research in Higher Education*, 30(6): 583-645.

Feldman, K.A. 1993, College Students' Views of Male and Female College Teachers – Part II: Evidence from Social Laboratory and Experiments. *Research in Higher Education,* 34 (2) 151-211.

Fernandez, M.A.M. 1997, Student and Faculty Gender in Ratings of University Teaching Quality. *Sex Roles: A Journal of Research*, 37: (n11-n12) 997 – 1003.

Gigliotti, R.J. 1987. Expectations, observations, and violations: Comparing their effects on course ratings. *Research in Higher Education*, 26: 401-415.

Gigliotti, R.J., & Buchtel, F.S. 1990. Attritional bias and course evaluations. *Journal of Educational Psychology*, 82: 341-351.

Gigliotti, R.J., & Seacrest, S.E. 1988. Academic success expectancy: The interplay of gender, situation, and meaning. *Research in Higher Education*, 29: 281-297.

Greenwald, A.G. & Gillmore, G.M. 1997a. Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52(11): 1209-1217.

Greenwald, A.G. & Gillmore, G.M. 1997b. No pain, no gain? The importance of measuring course workload in student ratings of instruction. *Journal of Educational Psychology*, 89(4): 743-752.

Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C. 1998: *Multivariate Data Analysis*, 5th ed. New Jersey, Prentice Hall, 603-604.

Haladyna, T. & Hess, R.K. 1994. The detection and correction of bias in student ratings of instruction. *Research in Higher Education*, 35(6): 669-687.

Hancock, G.R., Shannon, D.M, & Trentham, L.L., 1993. Student and teacher Gender in Rating of University Faculty: Results from Five Colleges of Study. *Journal of Personnel Evaluation in Education*, 6: 235-248.

Hatfield, L. & Kohn, J.W. 2003. Attribution Theory Reveals Grade-Leniency/Stringency Effects In Student Ratings Of Faculty, *Academy of Educational Leadership Journal*, Vol. 7, No.2.: 1-14.

Hatfield, L. & Kohn, J. W., 2004, Student Ratings of Faculty: Back to Square One - Integrating Theoretical Perspectives Using Structural Equation Modeling, *Academy of Educational Leadership Journal* , Vol. 1. No 10, 29-46

Holmes, D.S. 1972. Effects of grades and disconfirmed grade expectancies on students' evaluations of their instructor. *Journal of Educational Psychology*, 63(2): 130-133.

Howard, G.S. & Maxwell, S.E. 1980. Correlation between student satisfaction and grades: A case of mistaken causation? *Journal of Educational Psychology*, 72(6): 810-820.

Howard, G.S. & Maxwell, S.E. 1982. Do grades contaminate student evaluations of instruction? *Research in Higher Education*, 16: 175-188.

Howard, G.S., Conway, C.G. & Maxwell, S.E. 1985. Construct validity of measures of college teaching effectiveness. *Journal of Educational Psychology*, 77(2): 187-196.

Hoyt, D.P. 1973. Measurement of instructional effectiveness. *Research in Higher Education*, 1: 367-378.

Kennedy, W.R. 1975. Grades expected and grades received: Their relationship to students' evaluations of faculty performance. *Journal of Educational Psychology*, 67: 109-115.

Kline, R.B. 1998. *Principles and Practices of Structural Equation Modeling*. New York: Gilford Press.

Kohn, J.W. & Hatfield, L. 2001. Student Ratings of Faculty and Motivational Bias—A Structural Equation Approach, *Academy of Educational Leadership Journal*, 5(1):65-74.

Marsh, H.W. 1980. The influence of student, course, and instructor characteristics on evaluations of university teaching. *American Educational Research Journal*, 17: 219-237.

Marsh, H.W. 1984. Students' Evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76(5): 707-754.

Marsh, H.W. 1986. Self-serving effect (bias?) in academic attributions: Its relation to academic achievement and self-concept. *Journal of Educational Psychology*, 78:190-200.

Marsh, H.W. 1991. Multidimensional Students' Evaluations of Teaching Effectiveness: A test of Alternative higher-Order Structures. *Journal of Educational Psychology,* 83: 285-296.

Marsh, H.W. & Duncan, M. 1992. Students' evaluations of university teaching: A multidimensional perspective. In J.C. Smart (Ed.) *Higher education: Handbook of theory and research*, 8: 143-233. New York: Agaton.

Marsh, H.W. & Hocevar, D. 1985. Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin*, 97(3): 562-582.

MacCallum, R.C., Roznowski, M. & Necowitz, L.B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111(3), 490-504.

Marsh, H.W. & Roche, L.A. 1997. Making students' evaluations of teaching effectiveness effective. *American Psychologist*, 52(11): 1187-1197.

McHugh, M.C., Fisher, J.E. & Frieze, I.H. 1982. Effect of situational factors on the self-attributions of females and males. *Sex Roles*, 8:389-396.

McKeachie, W.J. 1997. Student ratings – The validity of use. *American Psychologist*, 52(11): 1218-1225.

Miller, D.C. 1991 *Handbook of Research Design and Social Measurement*. Newbury Park, California: Sage Publications, Inc.

Owie, I. 1985. Incongruence between expected and obtained grades and students' ratings of the instructor. *Journal of Instructional Psychology*, 12:196-199.

Powell, R.W. 1977. Grades, learning, and student evaluation of instruction. *Research in Higher Education*, 7: 193-205.

Ross, M. & Fletcher, G.J.O. 1985. Attribution and social perception. In G. Lindzey & E. Aronson (Eds.), *Handbook of Social Psychology* (vol. 2, pp. 73-122). New York: Random House.

Simon, J.G. & Feather, N.T. 1973. Causal attribution for success and failure at university examinations. *Journal of Educational Psychology*, 64: 46-56.

Stumpf, S.A. & Freedman, R.D. 1979. Expected grade covariation with student ratings of instruction: Individual versus class effects. *Journal of Educational Psychology*, 71: 293-302.

Tieman,C.R. & Rankin-Ullock,B. 1985. Student evaluations of Teachers. *Teaching Sociology*, 12, 177-191