THE ROLE OF VISUAL-SPATIAL APTITUDE IN ACCOUNTING COURSEWORK

by

Dianna Ross Coker

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

in Business

with a major in Accounting

APPROVED: Robert M. Brown, Chair

Larry N. Killough

Dirk Baldwin

Joseph J. Franchina

Kusum Singh

February 22, 1993

Blacksburg, Virginia

THE ROLE OF VISUAL-SPATIAL APTITUDE IN ACCOUNTING COURSEWORK

by Dianna Ross Coker

Chairman, Robert M. Brown

(ABSTRACT)

Accounting education research has explained some variation in student performance by aptitude, attitude, and experience variables, as well as gender. The unexplained portion of variance, however, suggests the existence of unidentified variables. This study examines the relationship of visual-spatial aptitude (VSA) to student completion and continuing behaviors and to performance in four accounting courses. VSA is a group of cognitive abilities which facilitate building mental representations and solving problems and which are positively related to performance in mathematics and science courses.

This study hypothesizes that high VSA students will complete Accounting Principles I at a higher rate, continue to Principles II at a higher rate, and perform better in Principles I than will low VSA students. Also hypothesized are gender differences and course differences in the relationship between VSA and performance.

Subjects are students tracked in accounting courses for three semesters. Independent variables include gender, prior bookkeeping coursework, and major, as well as SAT scores, GPA, and scores on two VSA tests--the MAP Planning Test (MAP) and the Mental Rotations Test (MRT). MAP and MRT measures include

the number right, the number wrong, and the percentage right. Dependent variables include student completion and continuing status as well as performance scores in each of four courses.

Results indicate that high VSA subjects have a higher completion rate than do low VSA subjects and that completers of Principles I have higher VSA than do droppers. Also, continuers to Principles II have higher VSA than do noncontinuers.

Results indicate relationships between VSA and Principles I scores. Subjects with high MRT percentages score higher on exams and lower on homework/quizzes than do those with low MRT percentages. Subjects with few MRT wrong have higher exam scores than do those with more MRT wrong.
MAP is related only to Principles II and Intermediate, while MRT is related only to Principles I, II, and Cost. Relationships of VSA to exams are positive and frequent. Relationships to homework/quizzes are negative and less frequent. Computerized practice set scores are rarely related to VSA.

In separate analyses of students taking Principles I, VSA is related to homework and exams for females and only to exams for males. For the smaller sample of students continuing to Principles II, models which contain general aptitude covariates indicate that VSA is only related to female performance and only in Principles I.

Acknowledgements

I am very grateful to the following people for their significant contributions to this dissertation.

Elena Vergara, Jose Cardenas, and Larry Crumbley, for insight, advice, and experience which led me to pursue this degree.

Bernadette Ruf, whose work introduced me to the topic.

Lorraine McClenny, Mary Harsh, and Greg Kane, for their valuable participation in the Disgang proposal workshops.

Rosalind Cranor and her teaching assistants, for their gracious assistance in gathering data.

My department head, Wayne Leininger, for his support throughout my years at Virginia Tech.

Members of my committee--Larry Killough, Dirk Baldwin, Joseph Franchina, and Kusum Singh--who were each responsive, supportive, and a pleasure to work with.

My committee chair--Robert M. Brown--for his abiding interest in this research topic, his invaluable guidance on technical topics, and his gentle encouragement through it all.

And especially my husband Jerry, for his initial, daily, and long-term encouragement to undertake and complete this research and this degree.

Table of Contents

List of Tables

Chapter 1 -- Introduction

Learning accounting requires the novice student to build a mental representation of a hierarchical information system in which sequential and parallel processes transform transaction data into financial statements. Manual bookkeeping can be seen as a sequence of steps from transactions to journal to ledger to trial balance to financial statements. Computerized bookkeeping is experienced more as parallel processing, since the original data are viewed from different perspectives to create reports which may be account histories, a trial balance, or financial statements.

Introductory accounting students may be able to reproduce the sequential processing steps but be unable to quickly judge the effects of transactions on financial statements without methodical mental progress through the sequence of the accounting cycle. Progressing from a mechanical processing ability to a more facile understanding of the flow of information would seem to rely on the parallel processing which is characteristic of spatial visualization.

The primary research question of this study concerns the relationship of student performance in accounting education to visual-spatial aptitude (VSA). VSA is aptitude in reasoning with abstract visual images (Lohman and Kyllonen 1983) and skill in representing, transforming, generating, and recalling symbolic, nonlinguistic information (Linn and Petersen 1985). VSA has been shown to affect learning in other disciplines such as mathematics, chemistry and physics. Performance in accounting education has been shown to be affected by individual attributes other than VSA, such as mathematical ability. The tasks involved in learning accounting suggest that VSA may also be an important determinant of success in accounting coursework.

Aptitude theory seeks to explain the relative match of aptitudes and tasks as a necessary first step toward improving two basic aspects of learning: aptitudes and instructional treatments. The improvement of learning from instruction can be approached through designing treatments supporting (a) inaptitude circumvention and/or (b) aptitude training once the aptitude's role in task performance is known. There is no published evidence to date which establishes the role of VSA in learning accounting. The combination of accounting education research literature and VSA research literature, however, makes a strong case for the investigation of the relationship between the two.

This report outlines theoretical and empirical support for examining the role played by VSA in student performance in accounting courses, discusses hypotheses generated from relevant theory, and presents results of research designed to test these hypotheses.

Chapter 2 -- Literature Review

The review of research literature begins with accounting education research and proceeds to studies on visual-spatial aptitude (VSA) and its role in learning and behavior. This organization first places the current research in the context of related research which has attempted to identify aptitude and other factors which are related to success in introductory, upper-level, and graduate financial and managerial accounting coursework. There is a great deal of performance variation yet to be explained, however.

Accordingly, this chapter reviews literature which establishes VSA as an aptitude related to coursework in other disciplines, as well as in nonacademic endeavors. Chapter sections are as follows: (1) accounting education and explanatory student attributes, (2) the nature of VSA, (3) the relationship of VSA to academic and work performance, (4) gender differences in VSA, (5) gender differences in the VSA-performance relationship, and (6) the role of VSA in accounting education.

<u>Review of related accounting education literature</u>. Accounting education research has investigated relationships between various student attributes and student attainment measures in undergraduate and graduate accounting courses. The instructional approach in most of these studies has been

traditional.[1]

Educational performance in accounting coursework is generally measured by grades on homework, quizzes, practice sets, and particularly exams. The behavioral variables in accounting education research have generally focused on individual exam scores or exam averages (because exams are the major determinants of the semester grade), course total scores, and course completion rates.

The research interest in grades seems rather straightforward. Accounting education researchers are responsible for providing a service, measurable by the knowledge gained by the student, for which the grade earned is a proxy. Also, research has indicated a relationship between grade point average (GPA) in accounting courses and students'

---

[1] Traditional instruction, at the introductory level, is typically lecture-centered, instructor-controlled, and problem-supplemented. There is rarely group project work or self-paced material coverage. Students are often more passive than active participants in the classroom due to the lecture orientation. Course-related student activity outside of class primarily consists of reading the text, working problems, and preparing a practice set, sometimes computerized.

Traditional instruction provides little intended differential treatment of students with the exception of providing different courses for business majors and nonmajors and for undergraduate and graduate students. Traditional instruction has been the most common treatment in which individual attributes have been studied in introductory accounting. Research in courses beyond introductory accounting may very well involve more heterogenous approaches, although most studies of upper-level courses do not provide enough course description to make that judgement.

subsequent performance on the Certified Public Accounting examination (Dunn and Hall 1984). Research interest in course completion rates is of practical importance also. Understanding the factors related to the behavior of dropping a class could be helpful in counseling, course design, and instructional decisions.

Independent variables used to explain drop rates and performance measures have included various student aptitude, experience, and attitude variables, as well as gender. The remainder of this section is organized as follows: (a) a discussion of student variables used in accounting education research for each general variable category, (b) a brief overview of findings for each area, and (c) detailed summaries of studies in this research stream.

Independent variables in accounting education research

General ability has most often been operationalized as scores on exams such as SAT, ACT and GMAT for entry into a college program (general math and verbal ability as evidenced by a single test) or as actual grades earned (general academic ability as evidenced by achievement). General ability achievement measures have included high school rank, high school and college GPA, and GPA in specific subsets of high school and college courses. GPA reflects the student's ability to apply aptitudes in the face of course requirements and outside distractions.

One analytic approach has directly examined the relationship between the general ability measure and performance. A second approach has tested for equality of mean general ability between groups of interest. These general ability measures have been used as proxies for learning capacity (Schroeder 1986) and quality of student (Bergin 1983, Elikai and Baker 1987).

Studies have examined experience variables such as college major, having taken a previous bookkeeping/ accounting course (generally in high school), the extent of previous accounting coursework, the extent of previous math coursework, related work experience, and frequency of reading business publications. Also included in the experience category are pre-test measures of accounting knowledge (gained from previous accounting exposure) including grade in last previous accounting course and score on a test from the AICPA Aptitude Test Bank.

The role of student gender has been studied as well as the interaction of student gender with instructor gender. Proxies for attitude have been measures thought to capture learning motivation, such as academic major or the number of quizzes a student elects to take.

Overview of Related Research

Student drop rates and aptitude/experience factors. Students who dropped an undergraduate introductory course had lower SAT (general ability, single test), high school GPA (general ability, achievement) and previous accounting coursework (experience) than did those who completed the course (Eskew and Faley 1988). One study found that the dropout rate for subjects with previous exposure to bookkeeping was 10%, while the rate for subjects without such exposure was 26% (Bergin 1983). Another study found that drop-out rates were lower in both Introductory and Intermediate accounting for those students with high school bookkeeping (7% and 16%) than for those without high school bookkeeping (22% and 30%) (Schroeder 1985). Another study identified that the timing of dropping out differed between those with and those without high school bookkeeping.

In another study, nonaccounting majors with no previous accounting coursework dropped from undergraduate introductory accounting more frequently than did nonaccounting majors with varying amounts of previous accounting coursework. Accounting majors with differing amounts of previous accounting coursework did not have different drop rates (Schroeder 1986).

Student performance and aptitude measures. Grades in introductory accounting have been shown to be positively related to SAT total scores (Eskew and Faley 1988), SAT Math

(McNeill and Collins 1975), ACT total (Doran, Bouillon, and Smith 1991; Schroeder 1986), college GPA (Doran, Bouillon, and Smith 1991; Eskew and Faley 1988; McNeill and Collins 1975), high school GPA (Eskew and Faley 1988), and rank in high school (Schroeder 1986).

Performance in upper-level undergraduate accounting courses is positively related to college GPA (Hicks and Richardson 1984; Ott, Deines and Donnelly 1987; Doran, Bouillon, and Smith 1991), ACT and grade in last previous accounting course (Doran, Bouillon, and Smith 1991). Eckel and Johnson (1983) found GPA and ACT score significant in discriminant analysis predicting student success in (defined as graduation from) an accounting undergraduate program.

Performance in graduate-level introductory financial accounting has been shown to be positively related to undergraduate GPA (Moses 1986), unrelated to undergraduate GPA if GMAT is controlled for (Ault and Carver 1987), and unrelated to GMAT once previous accounting knowledge is controlled for (Ault and Carver 1987).

<u>Student performance and attitude measures</u>. Discipline-specific attitude, as proxied by major, was found significantly related to exam scores in an introductory accounting course; students intending to become accounting majors outperformed those who did not (Schroeder 1986). Attitude, as measured by number of quizzes taken, was also

found to be positively related to performance on exams in introductory accounting (Eskew and Faley 1988), explaining more variance than was explained by college GPA and having had a previous bookkeeping course.

Student performance and experience factors. Positive experiential determinants of undergraduate introductory accounting performance include previous accounting coursework (Bergin 1983; Doran, Bouillon, and Smith 1991; Eskew and Faley 1988; Friedlob and Cosenza 1979; Schroeder 1985) and number of college statistics and math courses (Eskew and Faley 1988), but neither age (Edmonds and Alford 1989) nor completed college hours (Eskew & Faley 1988). Accounting majors outperform nonmajors (Doran, Bouillon, and Smith 1991).

Undergraduate upper-level accounting performance is related to major (accounting majors outperform nonmajors) and to grade in last previous accounting course (Doran, Bouillon, and Smith 1991; Hicks and Richardson 1984) as well as pre-test of accounting knowledge (Hicks and Richardson 1984). No relationship, however, was found between performance in Intermediate and the existence or extent of high school bookkeeping (Schroeder 1985).

Demonstrated experiential determinants of graduate-level introductory accounting performance include major (undergraduate business majors outperform nonmajors) (Ault and Carver 1987) and the existence of, but not the extent of, previous accounting coursework (Canlar 1986). However, the positive relationship early in the semester between undergraduate accounting coursework and graduate performance disappeared for measures later in the semester.

Student performance and the gender factor. Gender studies have produced conflicting results. Edmonds and Alford (1989) found no gender/performance relationship in introductory accounting when controlling for score on an AICPA Aptitude Test. Tyson (1989) found that females in undergraduate introductory accounting had SAT scores equivalent to those of males, had higher GPA than did males and demonstrated higher performance than did males. However, when GPA was taken into account, the gender/ performance difference disappeared. Doran, Bouillon, and Smith (1991) found that even controlling for GPA and ACT, males had grades superior to those of females in the first semester of introductory accounting (Principles I), but that male and female performances in Principles II were equivalent. Buckless, Lipe, and Ravenscroft (1991) reported mixed results using SAT as a covariate in analysis; male students in introductory accounting outperformed females at only one of three universities studied.

Mutchler, Turner, and Williams (1987) found that females outperformed males in upper-level accounting courses and that both females and males had higher scores when taught by instructors of the same gender (same-gender effect), but the study failed to control for either general ability or experience. Lipe (1989), who also controlled neither general ability nor experience, found no gender/ performance differences in an upper-level accounting course (managerial accounting) but did find a partial same-gender effect for male instructors only. Buckless, Lipe, and Ravenscroft (1991), controlling for SAT, found that intermediate accounting males outperformed females at a university where both male and female student scores were higher under male instructors.

Details of Related Studies

In the overview above, findings were presented within categories (general aptitude, attitude and experience findings, and gender differences). The details of these studies, however, are organized differently because many studies examined variables from more than one category. The particulars of methodology and analysis are reviewed below for studies presented in chronological order within academic level (principles, upper-level, graduate) of coursework studied. Details of those studies which examine more than one academic level are reviewed at the lowest academic level investigated. There is one exception to this organization; studies whose primary hypothesis concerns gender differences are presented in chronological order after the discussion of student determinants of performance in graduate-level accounting courses.

Undergraduate principles of accounting. McNeill and Collins (1975) used regression to examine the comparative role of aptitude in alternate treatments (traditional and self-study) in an introductory accounting course. Although the focus was on treatment differences, their results are discussed here in light of aptitude findings for the lecture-only treatment. The aptitude measures were personality measures (Edwards Personal Preference Scales) as well as conventional aptitude variables (SAT and GPA).

Subjects were enrolled in sections of introductory accounting which involved either self-study or lecture-only treatments on several chapters. For the lecture-only subjects, the two highest regression coefficients were for GPA and SAT-Math, both positive, confirming the relationship between general academic aptitude and performance in accounting. Although 59% of the grade variance was explained by the regression with the two conventional aptitude measures and three personality measures, it is unclear how much variance would have been explained by GPA and SAT-Math alone.

Friedlob and Cosenza (1979) examined the relationship between experience factors and performance in college introductory accounting. Chi-square results indicated a relationship between the number of semesters of high school bookkeeping (HSB) and performance in first quarter but not second quarter grades. A separate Chi-square found a relationship between prior accounting work experience and first quarter but not second quarter grades. First quarter college grades were also found to be strongly related to second quarter grades.

Baldwin and Howe (1982) also compared the performance in college introductory accounting of students without HSB to those with HSB. Repeated measure ANCOVA used as covariates the scores on vocabulary and quantitative skills diagnostic tests. Although the dichotomous HSB variable was not significant, the interaction of examination with HSB was marginally so, at $p < .084$. Although no tests of significance were provided, it was reported that students with HSB outperformed others on the first exam, performed comparably on the second and third exams, and performed less well than those without HSB on the final exam.

Baldwin and Howe also studied the relationship of HSB to drop-out rates and were unable to reject a hypothesis of no relation. Chi-square did find, however, that the timing of dropping out differed between the two groups. Those subjects with HSB dropped out as follows after each of three examinations: first (36%), second (50%), and third (14%), while 60%, 34% and 6% of the subjects without HSB dropped out after the first, second and third exam.

Bergin (1983) used Mann-Whitney, a nonparametric test of independence of groups, to examine the relationship of HSB to completion rates and to exam performance in introductory accounting. Subtotals on conceptual versus problem-solving exam questions were also analyzed. Subjects were enrolled in Fall 1980 in four sections of introductory financial accounting taught by one instructor. On the basis of self-report, subjects were categorized as having (43%) or not having (57%) HSB, without regard to the number of semesters of HSB. Nor was any information gathered as to the elapsed time period since taking HSB.

A Mann-Whitney test found GPA (gathered from Registrar records) not to differ between the two groups; accordingly, GPA was not included in subsequent analysis. For the students completing the course, Mann-Whitney tests were applied to three exams and a comprehensive final and to the three content-differentiated parts of each exam. On the first exam, subjects with HSB outscored those without HSB on the exam as a whole and on each of the parts (concept multiple choice, computational multiple choice, and problem-solving). The results for the second exam were similar and significant, but the difference between groups was narrowed. The trend had disappeared or reversed by the third exam; the only differences were for the exam as a whole and the problem-solving section. Subjects without HSB scored significantly higher on both. Unfortunately, the instructor announced experimental results before the final exam confounding the final exam results and the results of the semester as a whole, for each of which no difference found between groups of interest. The author suggests that any early advantage held by students with HSB is negated by subsequent complacency and under-achievement.

Schroeder (1985) found that students with two years of HSB outperformed both those with only one year and those without HSB in the first quarter, but not the second, of introductory accounting. No relationship was found, however, between the existence or extent of HSB and grades in Intermediate.

Schroeder (1986) examined aptitude, attitude and experience variables for relationships with the rate of completing a college accounting course and performance in that college course. Subjects were students enrolled in 18 sections of introductory financial accounting.

The relation of HSB to probability of course completion was examined with Chi-square on three levels of HSB: none, one year or less, and more than one year. Significant completion rate differences were identified between the subjects with no HSB and those with one year or less, and between those with no HSB and those with more than one year. Drop rates were similar for those with one year or less and those with more than one year. Analysis of academic major and HSB indicated that nonaccounting majors with HSB were more likely to complete the course than nonaccounting majors without HSB.

The analysis of the relation of HSB to course grades used exam scores which had been standardized between instructors by transforming scores into a percentile rank equal to the percentage of students taking the same exam who received either the same or a lower score. Although Mann-Whitney rank sum tests indicated that students with one year or less HSB outperformed students with no HSB on the first exam, they also showed that students with more than one year HSB outperformed both other groups consistently in all exams in the course.

Repeated measure ANCOVA controlled for effects other than HSB levels and found results consistent with Mann-Whitney. In addition to the experience variable HSB, independent variables included learning capacity (proxied by ACT comprehensive score and high school class rank), availability of study time (asked directly and proxied by academic course load and employment hours), related work experience, and intended academic major. Major was thought to capture course-specific motivation, since students would be more motivated to work in a course in their area. All measures were self-reported except ACT and high school ranking, which were obtained from university records. In addition to HSB, the following covariates were related to course performance: ACT, high school class rank, and major.

Accounting majors, students with high ACT scores and those with higher class rank in high school outperformed nonaccounting majors and students with lower ACT scores and lower high school class rank. Nonsignificant covariates included availability of study time, academic course load, and employment work hours. Schroeder concluded that the effect of the variable of interest, HSB, is dependent upon the amount of HSB. While one year or less may be related to superior performance on early exams, consistently superior performance was related to more than one year of HSB.

Eskew and Faley (1988) used regression to examine the role of aptitude, attitude and experience factors in determining drop rates and performance in introductory accounting. Subjects were 425 students enrolled in one large section of introductory accounting at Purdue University in Fall 1983. Student assignment to the experimental section was accomplished by computer and was considered random. Of the original sample, 73 students dropped and 352 completed the course in that semester. Independent variables included SAT scores (the authors used the sum of SAT-Math and SAT-Verbal, since correlation between these measures was .5495), high school grades (in math and English), college GPA, previous bookkeeping/ accounting coursework (dichotomous: yes or no), previous related coursework (hours of math and statistics), number of completed college course hours, and number of quizzes. This last variable was included as a proxy for motivation, since the quizzes earned few points (relative to the course total) and were considered a voluntary activity.

Hotelling's T^2 found a significant difference in the vector of independent variables for students who dropped and those who completed the course. Individual t-tests found that students who dropped had lower values on SAT scores and high school grades (math and English) than did students completing the course. Also, the percentage of dropping students who had previous bookkeeping/accounting coursework was less than the percentage of students completing the course who had no such background work.

The same vector of independent variables was used in a regression to explain course performance variance. The dependent measure was the sum of three midterm exams and a final exam. R^2 for the multiple regression was 54%, and all independent variables were significant except for the number of completed college hours. Coefficients of partial determination indicated that variables ranked from highest to lowest marginal contribution were: SAT scores (.2034), number of quizzes (.1217), high school grades (math and English) (.0699), HSB (.0425), college GPA (.0205), and previous related college coursework (math and statistics) (.0178). Regressions on individual exams indicated that previous bookkeeping/accounting coursework had a positive effect on all exams, not just earlier exams as had been indicated by earlier research (Bergin 1983, Baldwin and Howe 1982). The authors concluded that readily available aptitude measures (SAT, high school grades,and GPA), attitude indicators (number of quizzes), and experience variables (HSB and related college coursework) can explain more than half of the variance in Principles I performance, and that many of these same variables differed between course droppers and completers.

Edmonds and Alford (1989) used ANCOVA to examine the relationship between exam performance in accounting principles between two treatment groups (subjects taught through either a high-complexity or low-complexity model). This study is included here, not for its analysis of treatment differences, but for its examination of three covariates: age, gender, and an aptitude/knowledge pre-test (score on AICPA Aptitude Test Form Z). Subjects were 68 students enrolled in the first semester of financial accounting principles at a large urban state university. The dependent measure of interest was the first exam score following introduction of the accounting cycle (looking for immediate treatment effects under alternate teaching approaches) and the remaining exams (looking for long-term effects). Treatment effects were found on the first exam but not on the subsequent exams. Although neither age nor gender were explanatory for exam scores, the aptitude measure was significantly related to all exams for both treatment groups at $p < .0223$.

Doran, Bouillon and Smith (1991) also used regression to examine the association between performance in the first two semesters of accounting and aptitude, experience and gender, with an eye toward providing guidance to more efficient enrollment policies. Following Eskew and Faley (1988), they examined the relation between the sum of exams and GPA, ACT, HSB, and academic major. Following gender research (Mutchler, Turner and Williams 1987; Lipe 1989), they also analyzed gender and the gender match between student and instructor. They extended the research stream, however, by also examining the relation between the grade in the first accounting course with the grade in the second and by analyzing the association between the first exam and remaining exams.

Subjects were enrolled in 1987 in Principles I and Principles II. All independent variables were from self-report survey (including three variables subsequently excluded from the model due to nonsignificance: expected hours of enrollment, work, and study). Subjects reported their ACT scores as less than 20, more than 20 less than 24, more than 24 less than 28, more than 28 and their cumulative GPA in half-point ranges from less than 2.0 to greater than 3.5). Major was self-reported as either accounting or nonaccounting. The authors' choice to examine HSB only as a dichotomous variable only is more understandable if one remembers that

Eskew and Faley's HSB results were not published until 1988, a year after Doran, Bouillon and Smith collected their data.

Regression (model 1) for Principles I (adjusted $R^2 = .30$) found significant positive relationships of exam sum with GPA, ACT, accounting major, gender and HSB (listed in order of standardized beta coefficients, highest to lowest). Male students outscored female students in this sample. A dummy variable which corresponded to sameness of student and instructor gender was not significant. The smaller R^2, relative to Eskew and Faley (1988), was thought to be attributable to the ordinal aptitude measures, the exclusion of any effort measure, and the noise from including subjects from numerous sections, instructors and times of day.

A higher adjusted R^2 (.52) was found for a regression (model 2) in which the dependent measure was the sum of all exams except the first exam, and the independent measures were all those used in regression model 1 with the addition of the score on the first exam. The first exam exhibited the strongest relationship to the remaining exams, and accounting major and HSB were no longer significant. GPA, ACT, and gender were still significant.

Model 1 regression for Principles II (adjusted $R^2 = .55$) included the Principles I total grade in addition to GPA. The Principles II exam sum was significantly and positively related to Principles I total, GPA, and ACT (ranked highest beta coefficient to lowest) and negatively related to HSB. The exam sum was also marginally and positively related to accounting major. Gender was no longer significant, and student-instructor gender match was still not significant. Model 2 regression ($R^2 = .59$) indicated a strong relationship between the first Principles II exam and the remaining exams. As in the regression models related to Principles I, major and HSB were no longer significant. Variables ranked from most to least explanatory were: Principles II exam 1, Principles I total, and GPA. ACT was no longer significant.

This study was not a longitudinal one, so the authors are not able to compare regression results for one sample of subjects in each course. It is difficult to know how to interpret the differences found between the Principles I and Principles II parameter estimates. The different relationships for HSB and gender may be attributable to sample differences or underlying differences in the course requirements. The authors are able to conclude, however, that aptitude (GPA, ACT), attitude (major), and experience (HSB) are related to Principles I performance. Even stronger relationships are identified for more discipline-specific aptitude indicators such as first exam score or grade in the previous accounting course. Whereas models without these discipline-specific indicators suggest that the experience of having a high school bookkeeping course is positively related to Principles I and negatively related to Principles II performance, models including such indicators suggest that HSB is no longer uniquely explanatory. For these models, aptitude indicators alone explain more than half of the variance in accounting grades.

Upper-level undergraduate accounting courses. Eckel and Johnson (1983) used discriminant analysis to study the variables related to upper-level accounting GPA for accounting students who graduated from one university in June 1979. The dependent measure was operationalized as GPA for the six upper-level required accounting courses: two intermediate financial, two intermediate managerial, one intermediate auditing, and one accounting elective. Students were classified into higher GPA (2.5 or more on a 4 point scale) and lower GPA (less than 2.5). For the prediction sample of 69 students[2], the discriminant analysis function examined the ACT Math and ACT Composite scores and eight grade-related variables: lower level accounting GPA (average in two accounting principles courses), GPAs in two lower-level required math courses and two required stat courses, grade in a lower-level management information systems course, grade in

---

[2] The prediction to holdout sample ratio was determined by sensitivity analysis of iterations of discriminant analysis with different assumptions of prior probabilities of high and low GPA. The ratio of 72% to 28% minimized type I errors (predicting a success as a failure).

last required English course, and sophomore GPA.

The discriminant function included as significant only two of the independent variables: lower level accounting GPA and ACT Math score. Use of this model correctly classified 85% of the holdout sample while maintaining a type I error rate of 5.3% and type 2 error rate of 37.5%. Acceptable classification rates were obtained with just two measures, ACT Math and lower-level accounting GPA, both available at the time decisions are made about admittance as an accounting major.

Hicks and Richardson (1984) used regression to explore the relationship of performance in Intermediate Accounting and two measures of accounting knowledge: the average of the grades earned in the first two prerequisite Principles courses and a researcher-designed pre-test. This pre-test covered the accounting cycle and was given at the end of the first week of class. Subjects were students enrolled in seven sections of Intermediate Accounting taught by four instructors. Six simple linear regressions were run each with a dependent variable of the course total or the course grade and an independent variable from pre-test, Principles GPA or freshman-sophomore GPA. All were significant, with R^2 ranging from .157 to .348. Multiple linear regressions were run alternately for course total or course grade and included pre-test, either Principles GPA or freshman-sophomore GPA, and instructor variables. All models were significant for pre-test and GPAs, with one instructor variable contributing to the explained variance. R^2 was in the high 40% for all four models. Unexplained grade variance, indicated by Root MSE, was reduced much more by Principles GPA than by freshman-sophomore GPA. The authors discussed no models involving both GPA measures.

Ott, Deines and Donnelly (1987) used aptitude measures in ANCOVA to examine the relationship between performance in Intermediate Accounting and a practice set which reviewed material taught in introductory financial accounting, the prerequisite course(s) for Intermediate Accounting. GPA was used as a covariate while testing for treatment effects, instructor-time effects, and preknowledge effects measured by grade in introductory financial accounting. Treatment affected performance, as did GPA, instructor/time, and preknowledge. GPA was the greatest contributor to performance.

<u>Graduate-level accounting courses</u>. Canlar (1986) used Mann-Whitney tests to study the relationship between performance in the first MBA-level financial accounting course and previous college-level exposure to accounting. Subjects were students from two evening sections of the graduate financial accounting course taught by one instructor during Fall 1983. These subjects were categorized as not having or

as having had previous accounting coursework, and then the latter group was subdivided into those with one and those with more than one such course. GMAT scores were not used in the hypothesis tests, but were used to establish equivalence of groups between those with and those without previous accounting coursework. The authors did not report, however, testing for GMAT equivalence between the subgroups of one course versus more than one course.

Dependent variables included each of three exams and the exam total. Mann-Whitney showed declining differences between groups for the first two exams and for the exam total (with higher scores earned by subjects with previous coursework), but no differences were found for the final exam. Tests between subgroups also indicated no difference in performance between those with one versus those with more than one previous college-level accounting course.

Ault and Carver (1987) used covariates of interest to the present study in regressions exploring the relationship between graduate-level performance in financial accounting and pedagogical format (weekly evening classes versus two weekend classes combined with independent study). This study, too, is included for its findings related to covariates, principally undergraduate GPA and GMAT scores (quantitative and verbal). Both GMAT-Verbal and GMAT-Quantitative were explanatory when examining the financial accounting post-test score, but not explanatory for the financial accounting gain score (post-test less pre-test) or the managerial post-test score. Neither GPA effects nor treatment effects were found for any of the three dependent measures. Of particular interest here, however, is the relationship found between the financial accounting post-test score and the managerial post-test. This was the only relationship identified in the regression for managerial post-test. This is comparable to using the Principles I grade as a covariate when examining the Principles II grade in an undergraduate study.

Moses (1987) used regression to examine the relationship between performance in graduate-level accounting and three sources of prior exposure to accounting and/or financial concepts and ideas: the number of undergraduate accounting courses, the years of accounting/ finance employment experience, and independent reading of business/financial publications. The reading variable was self-report (as were all independent measures) on a 5-point scale of frequency of business reading: never, rarely, occasionally, frequently, and daily. Subjects were 94 graduate students enrolled in four sections of first-term financial accounting taught by two instructors. All sections had common homework assignments, text, and exams. Mann-Whitney U-tests revealed no differences between subjects taught by different instructors on the dependent variable (number corresponding to letter grade for semester), on the three independent variables of interest mentioned above, and on GPA. Although regression coefficients were positive and significant for GPA, previous accounting employment, and business reading, this was not quite the case for number of accounting courses, for which a p-value of .054 was obtained. The amount of grade variance explained was only 18%.

The author also ran additional regressions on subsamples arrived at by dichotomizing the independent variables (with and without accounting coursework, with and without accounting work experience, and frequent and less frequent readers of business publications. One conclusion of relevance here was that two variables--accounting work and frequency of reading--appeared to be substitutions but to not behave additively, suggesting the reasonableness of including only one of these variables in models in related research.

<u>Examination of gender differences in accounting courses</u>. Mutchler, Turner and Williams (1987) reported results of a longitudinal study and a cross-sectional study which examined gender differences in accounting course performance. For both studies, the dependent variable was a standardized score for points earned. The score was (a) points earned less the mean of that section divided by (b) the standard deviation of that section.

Subjects in the longitudinal study were students in one male instructor's auditing course over an 18 year period. T-tests indicated that females outperformed males in 13 of those years. There were, however, no tests for gender equality on aptitude, attitude, or experience variables.

Subjects in the cross-sectional study were enrolled in an unreported number of sections of unspecified upper-division accounting courses. No aptitude, attitude, or experience variables were considered in the analysis. Analysis of variance (ANOVA) results indicated that females outscored males, that females performed better with females instructors, and that males performed better with males instructors. Also, grades for both males and females were higher with female instructors.

Lipe (1989) also used ANOVA to examine student gender effects on performance in an introductory managerial course and attempted to control for grading policy confounding by using the percentage of the possible points on common exams as the dependent measure. The author found an interaction between the student gender and instructor gender. For males instructors, males scored higher than females. For female instructors, there were no gender differences, unlike Mutchler et al. (1987).

Although the results of both studies suggest either gender differences and/or interactions between student and instructor gender, the studies lacked appropriate covariates.

A subsequent study attempted to remedy this deficiency with respect to aptitude.

Tyson (1989) used various statistical approaches to examine gender differences in how three intrinsic achievement components relate to performance in undergraduate introductory financial accounting. Subjects were students enrolled in a principles course who consented to experimenter access to SAT and GPA records. The primary independent measure was score on a personality test for which no reliability or validity numbers are provided. Of primary interest here, however, are the relationships of aptitude to gender which were reported. Although no gender differences were found for SAT-Math and SAT-Verbal, females had higher GPA and higher standardized score (dependent measure). When GPA was used as a covariate in ANCOVA, there was no gender difference in the dependent measure. Regression indicated that GPA and SAT-Math were both significantly related to the standardized score. With these covariates in the regression, however, gender was not explanatory. Lipe did not examine the possible interaction between student gender and instructor gender, nor did she control for instructor grading effects.

Buckless, Lipe and Ravenscroft (1991) attempted to synthesize and extend the work of Mutchler et al., Lipe, and Tyson by examining the gender/performance issue with academic aptitude covariates and with experimental design controls for individual instructor grading effects. Subjects were enrolled in the first semesters of Accounting Principles at three universities and enrolled in Intermediate Accounting at one of the universities.

Dependent measures were scores on final exams which, within each university, were commonly administered and computer-graded. Independent variables were student gender and instructor gender. SAT scores and high school GPA (both self-reported and in some cases from official records) were considered as possible covariates. Analysis of each measure against the official record available in at least one of the three schools indicated a self-reporting bias in GPA not present for SAT scores. Accordingly, SAT (the average of SAT-Math and SAT-Verbal) was included in ANCOVAs (official SAT used in two schools, self-reported in third).

Analyses were as follows: (a) ANOVA (student gender by instructor gender) in introductory accounting across all universities (exam scores at all three universities were standardized to a mean score = 0 and standard deviation = 1), (b) same model but with SAT covariate in ANCOVA, (c) similar ANOVA and ANCOVA for each university (using untransformed scores), and (d) ANOVA and ANCOVA of Intermediate scores.

In separate analysis of university number one, the gender difference was significant with ANOVA and with ANCOVA, with males outscoring females (inconsistent with Mutchler et al.);

no interactions were significant. For the second university, an interaction of student gender with instructor gender was identified, but with female students performing better with instructors of the opposite gender, unlike results reported by Mutchler et al. and Lipe. This effect disappeared when SAT was introduced into the ANCOVA, however. For university three, no variables were significant in ANOVA. SAT scores were significant in ANCOVAs for all universities.

ANOVA results showed that students with males instructors outscored those with female instructors, and (unlike Mutchler et al.) male students outscored female students. These results were also obtained in ANCOVA, in which SAT was also significant. The authors conclude that gender effects vary by university and differ from those reported by either Mutchler et al. or Lipe, and that some effects disappear when academic aptitude is controlled for, making some previous studies inconclusive. Interestingly, R^2 for ANOVA models ranged from .04, .01, and .01, whereas ANCOVA models were .14, .17 and .10, which are much lower than explanatory models in other studies already discussed.

<u>Accounting education research--summary</u>. Given significant but inconsistent gender findings, the studies reviewed above indicate the need for continued research into gender issues. It is also clear from studies reviewed earlier that many researchers have reported significant relationships

between accounting coursework performance and aptitude measures and some experience and attitude measures. The small amount of performance variation accounted for by these variables, however, indicates that much of the variation in accounting education performance is unexplained, indicating that important variables are missing from analyses to date. In their review of accounting education research, Rebele, Stout and Hassell (1991) recommend exploring the explanatory power of endemic factors other than general scholastic aptitude.

In looking for such variables, it is reasonable to look to those factors found to be significant in other related disciplines. SAT-Math has been shown to be related to accounting performance as well as to mathematics and science performance. This commonality leads one to examine other reported determinants of performance in mathematics and science coursework. Such other determinants may be possible explanatory variables in accounting education, as well. Visual spatial aptitude (VSA) is one such variable which has been shown to be related to science and math coursework as well as other adult behavior. Accordingly, the next sections review VSA research in support of developing VSA hypotheses related to accounting education (Chapter 3) and discussing research designed to test those hypotheses (Chapters 4 and 5).

Visual-Spatial Aptitude

VSA refers to a group of cognitive abilities which are distinct from verbal and mathematical abilities and which utilize spatial concepts. These multiple abilities are grouped as the general ability to use spatial information (Cooper and Mumaw 1985). Spatial information concerns spatial properties, such as shape, size, and position.

Spatial problems are those that have a significant amount of spatial information in the problem presentation or in the way an individual mentally represents the problem. Spatial problems are typically solved by generating a mental image of a two or three dimensional structure and then assessing its properties as is or after mentally transforming that representation by adding, deleting or altering properties (Carpenter and Just 1988).

VSA is thought to be situationally helpful; only some tasks are facilitated by these aptitudes and then only at certain stages in the task achievement (Snow and Lohman 1984). Examples of common activities which are facilitated by spatial aptitude are driving, anticipating the trajectory of a thrown object, and exploring an unfamiliar city.

Performance differences on spatial tasks are thought to be a function of differences in the accuracy and/or quality of mental representation, not just speed of processing, and speed differences may reflect differences in ability to form accurate mental representations (Mumaw and Pelligrino 1984). The quality of an individual's mental representation is not directly observable, however, and must be inferred from such overt behaviors as number of eye movements when examining an image (for instance, Carpenter and Just 1988). Individuals with high spatial abilities are believed to be better at generating, maintaining and coordinating information during spatial transformations (Carpenter and Just 1988), although it is not clear whether all high-VSA individuals encode in more detail or maintain more stability of information during transformations or both (Cooper and Mumaw 1985).

Because accuracy of spatial images and efficiency of spatial transformations are unobservable, aptitude tests measure accuracy and efficiency of performance on sets of spatial tasks thought to require VSA for processing. The impetus for development of VSA testing has been the desire to predict important real-world behaviors on the basis of VSA scores. Many VSA measures were developed in the first half of the century to aid in vocational selection decisions such as Air Force selection of pilot trainees and other aircraft personnel. In many technical training programs and some coursework, the predictive validity of VSA measures is greater than for measures of general or verbal ability (Cooper and Mumaw 1985).

VSA tests measure components of real-world skill not measured by other aptitude tests. Numerous VSA tests have been developed to measure different aspects of spatial ability. [See the International Directory of Spatial Tests (Eliot and Smith 1983) for descriptions, sample pages, and references for hundreds of such tests.] Measures of VSA have used tasks such as perception of the horizontal, mental rotation of objects, and embeddedness or the location of simple figures within complex figures (Linn and Petersen 1985). Some researchers have included, within broader published cognitive test banks, inventories of paper-and-pencil spatial tests corresponding to spatial factors; one such bank is the Kit of Factor-referenced Cognitive Tests (Ekstrom et al. 1976). Abilities required by paper-and-pencil tests are comparable to those required by performance tests which require the subject to physically manipulate concrete materials (Lohman 1988).

Factor analytic studies since 1925 have identified different taxonomies of spatial factors. McGee (1979) provides a detailed review of the first 50 years of this research stream. Although there is disagreement about the number and nature of spatial abilities, there are at least two separate factors. The two major factors have been called spatial visualization and spatial orientation, although the use of these terms has not been consistent.

Spatial visualization has been defined as the ability to mentally manipulate pictorially presented stimuli by a process which involves recognizing, retaining and recalling configurations in which there is movement of the figure or parts of the figure (McGee 1979) and as the ability to imagine how objects will appear when transformed in some way such as when rotated (Halpern 1986). Spatial relations (mental rotations) is a related factor and is measured by tasks which are similar to but less complex than those measured by visualization tasks (Kail and Pelligrino 1985). Spatial relations tasks require the matching of three-dimensional objects differentially rotated in space, requiring the individual to retain complex object information during the process of mental rotation (Lohman and Kyllonen 1983). The Cube Comparison Test (Ekstrom et al. 1976) and Vandenburg and Kuse's (1978) Mental Rotations Test are measures of spatial relations.

Spatial orientation has been defined as the ability to remain unconfused by changing orientations (McGee 1979) and as the ability to detect relationships among different stimuli and to perceive different patterns correctly (Halpern 1986). Spatial scanning is related to spatial orientation, concerns accuracy and speed in visually exploring a complicated spatial field or maze-type task, and is analogous to rapidly scanning a printed page for comprehension (Ekstrom, French, and Harman 1976). Spatial scanning has been measured using the Maze

Tracing Speed Test, Choosing a Path, or the Map Planning Test, all in the Kit of Factor-referenced Cognitive Tests, Ekstrom et al. (1976).

The Relationship of VSA to Academic and Work Performance

The ability to transform a mental representation while retaining enough information to correctly match it with an external referent gives high VSA individuals an advantage in many performance situations, reviewed below. Low-VSA individuals are unable to retain an accurate image of one stimulus while viewing additional stimuli (Mumaw and Pelligrino 1984, Lohman and Kyllonen 1983). VSA is thought to be involved in both general and scientific problem-solving, and in managing complex variable arrays and multidimensional networks of data (Baker and Belland 1988).

VSA has been shown to predict student performance on proportional reasoning tasks (Linn and Pulos 1983).

Poole and Stanley (1972) used factor analysis to examine the relationship between performance in engineering courses and VSA tests which loaded on a spatial visualization factor or a spatial orientation factor. Performance on engineering tasks which involved the interpretation or production of graphical information also reflected the two underlying VSA factors, as well as a factor interpreted as academic ability.

McGee (1979) and Meece et al. (1982) review earlier studies which demonstrated a relationship between VSA and performance in math and science courses. The following sections report details of more recent studies.

VSA and academic performance. Pribyl and Bodner (1987) reported that high VSA students perform better than low VSA students at early stages of problem-solving in organic chemistry. Subjects for that study were students in four courses at the West-Lafayette campus of Purdue University. The courses were analyzed separately because each was designed to serve a different student population (agriculture/health science majors, biology majors, chemistry/chemical engineering majors, or pharmacy/medicinal chemistry majors).

Dependent measures were exam scores. Independent measures were scores from two VSA tests--Purdue Visualization of Rotations (ROT) and a hidden figures test (Find-A-Shape-Puzzle or FASP)--and a total VSA score. This third variable (TSPAT) was the sum of standardized ROT and standardized FASP scores. Students were divided into three groups on each of these VSA variables for separate analysis. High scorers on each measure were those whose score was more than one-half standard deviation above the mean, while low scorers were those scoring lower than one-half standard deviation below the mean. No other aptitude measures were analyzed.

Separate two-way ANOVAs (VSA variable by gender) were performed for each of the 20 exams given among the four courses. Performance differences were found among ROT groups on six of the 20 exams, with only one exam showing a relationship to gender and another exam showing an interaction of gender and ROT. The nature of this interaction was not reported. Performance differences were found among FASP groups on nine exams, with gender related to only one exam. Performance differences were also identified among TSPAT groups on 12 exams; gender was related to two exams; there was no interaction found between TSPAT and gender. Results were strongest for the chemistry-major course; performance differences among TSPAT groups were found for every exam. Pearson product-moment correlation coefficients indicated that up to 15% of the variance in exam scores was attributable to VSA. Scheffe's tests indicated that, in 77% of the cases where ANOVA differences were indicated, high-VSA subjects received higher exam scores than low-VSA students. No information was provided about the other 23%.

After analysis of exam sub-parts, the authors concluded that the performance differences among VSA groups were attributable to the VSA-performance relationship on problem-solving items and/or problems requiring students to mentally manipulate two-dimensional representations of molecules. As striking as these results are, however, the study did not include general academic aptitude measures as covariates.

To ensure that the VSA measures reflected a factor distinct from general academic aptitude and test-taking ability, Carter, LaRussa and Bodner (1987) used the same categorical treatment of ROT, FASP, and TSPAT in ANOVAs and in ANCOVAs which included SAT scores. Subjects were students enrolled in Fall 1983 in multiple sections of three general chemistry courses, two for science/engineering majors taught on different schedules and a third for nursing/agriculture majors.

A positive relationship was found between VSA and exam performance. Results of two-way ANOVAs (VSA variable by gender) showed performance differences among VSA groups for all three VSA measures on all exams in both courses. These effects were also found for 24 of 35 sub-parts for ROT, 30 of 35 for FASP, and 32 of 35 for TSPAT. Scheffe's test indicated that high-VSA students outscored low-VSA students on all of the exams. ANCOVAs still found performance differences among TSPAT groups in 19 of the 32 tests where ANOVAs had done so. ANCOVAs were not reported for ROT or FASP. Only a few sub-part ANOVAs identified an interaction among VSA and gender, although the nature of this interaction was not reported.

Correlation analysis of sub-part scores with VSA indicated that correlations (of VSA measures with performance) were higher on verbally complex problems and items requiring problem-solving than on items answerable algorithmically or from memory. Although not discussed in any detail, authors reported that multiple regression using ROT, FASP and gender explained from .10 to .37 of exam score variance.

VSA and behavior other than coursework performance. VSA predicted novice adult performance in searching hierarchical computer files (Vicente 1987), female adult performance using a line text editor word processing package (Gomez, Egan, and Bowers 1986), and college students learning a pattern game (Wolf and Frey 1984-85).

Two of three studies on the relationship of VSA levels and completion rates in coursework or experimental tasks found a positive relation between VSA and course/task completion. Carter, LaRussa, and Bodner (1987) found no difference in VSA between students completing and dropping an introductory chemistry course. However, Pallrand and Seeber (1984) found that droppers has lower VSA than did completers of introductory physics. Further, adult female subjects who continued to advanced exercises in an experimental task had higher VSA than subjects not advancing to these exercises (Gomez, Egan, and Bowers 1986).

Gender Differences in VSA

Numerous studies have identified gender differences in VSA, with most results showing males had higher VSA scores than did females (McGee 1979). Maccoby and Jacklin (1974) reviewed 43 studies of analytic spatial ability and reported that males scored higher than did females on 22 studies, while females scored higher on only one. They also noted that female and male sample distributions of VSA generally overlap, often with greater variation in the male distribution.

Linn and Petersen's 1985 meta-analysis concluded that gender differences are large for mental rotation and medium for spatial perception. Hyde's 1981 meta-analysis found that gender explained about 1% of variation in verbal ability and in quantitative ability and 4% of variation in VSA. In terms of another measure of effect size, the statistic d, gender differences in VSA studies had a median value of .45, considered a medium effect size.

Some research indicates that males demonstrate more skill on VSA tasks than do females by early adolescence (Maccoby and Jacklin 1974, Pepin et al. 1985), while others find no such gender difference in VSA until late adolescence (Fennema and Sherman 1978). In a middle school study, discriminant analysis did not differentiate gender on the basis of VSA (Baker 1985). In high school, however, males outperformed females on a VSA task (Linn and Pulos 1983).

By college age VSA gender differences are consistently found. Multiple studies found that males outperform females on the Porteus Maze Task (McGee 1979) and in a spatial visualization task (Newcombe, Bandura, and Taylor 1983). Lord (1987) reported that male science majors had higher VSA than did female science majors, who had higher VSA scores than did male and female nonscience majors.

Linn and Petersen (1985) reported that males tend to outperform females on mental rotation tasks at any age where measurement is possible, with these differences possibly resulting from differential rate of rotation, differential efficiency in strategy application, differential use of analytic processes, or differential caution. Baker (1990) also reported that males outperform females on the Mental Rotations Test. Blough and Slavin (1987) reported that females had longer reaction times than did males on mental rotation tasks, which increased with the degree of rotation.

Research on environmental and biological explanations of gender differences in VSA has yielded limited and ambiguous support for various hypotheses. One environmental hypothesis concerns differential socialization and resulting differences in experiences. Biological hypotheses have explored whether VSA might be controlled by a recessive sex-linked gene, by differential brain lateralization, and/or by hormone differences (Benbow and Benbow 1984, Vandenberg and Kuse 1979). Meta-analysis suggests that the magnitude of gender differences in spatial ability has been decreasing over time, suggesting an environmental/experiential explanation for at least part of the VSA gender difference (Baenninger and Newcombe 1989).

Gender Differences in the VSA-Performance Relationship

In addition to the difference in VSA scores between females and males, gender has been found to be a moderating variable in the relationship between performance and VSA. This is reflected in interactions between VSA and gender, such as those reported by Pribyl and Bodner (1987) and Carter, LaRussa and Bodner (1987), referenced above.

Numerous other studies have examined gender differences in the relationship of VSA with performance on other aptitude tests and with academic performance. Baker (1985) reported that in middle school science, VSA was differentially related to course grade for males and females. Males who made A or B had higher spatial scores than did males who made C or D and females who made A, B, C, or D.

Meece et al. (1982) reviewed studies in which gender differences in math achievement scores were significantly reduced or eliminated when spatial skills were partialed out. These findings are consistent with the hypothesis that gender differences in VSA contribute to gender differences in math achievement.

Battista (1990) found that gender was a moderating variable when examining the relationship of high school geometry achievement to VSA, logical reasoning, and the discrepancy between VSA and logical reasoning scores. Males scores were higher than were female scores on spatial visualization, while female scores were higher than were male scores on geometry achievement. No gender differences were found for measures of logical reasoning. In stepwise regression for females, VSA accounted for 37% of the variance in geometry achievement. For males, 26% of variance in geometry achievement was accounted for by logical reasoning and by the discrepancy between reasoning scores and VSA scores, but not by VSA.

Brown (1991) also found that VSA predicted high school geometry achievement for females, while logical reasoning and short-term memory span predicted geometry achievement for males. Males scores were higher than those for females on measures of speeded rotations and spatial orientation, while females had higher scores than did males on short-term memory span and geometry achievement.

A VSA-gender effect was found for performance in college introductory chemistry for science & engineering majors in one class (male scoring higher in almost every dependent variable) but not in a second class (Carter, LaRussa, and Bodner 1987). VSA was more highly correlated with high school math and geometry credits for college women than for college men (Lunneborg and Lunneborg 1984).

Linn and Petersen (1985) suggest that adult males and females may differ in the processes they select for solving some spatial problems. For instance, performance on a paper folding task was correlated with VSA for males but with verbal ability for females (Kyllonen, Lohman, and Snow 1984).

Although various subfactors of VSA are assumed to be stable in adults in the absence of intervention (e.g., Pribyl and Bodner 1987), there is empirical support for the development of VSA in adults through directed training. The majority of these studies reviewed by McGee (1979) have shown no gender difference in training effectiveness. However, Lord (1987) found post-training performance gains greater for females, who were lower in VSA before training, than for males. Also, in meta-analysis of effect sizes across studies, Baenninger and Newcombe (1989) found no gender differences in improvement after VSA training.

Lohman and Kyllonen (1983) did not examine gender differences but did suggest that the effectiveness of interventions may rely on an individual's abilities other than VSA. For instance, they recommend that VSA training emphasize visualization for low-VSA, low-verbal students and a nonvisual strategy for low-VSA high-verbal students. This recommendation implies an interaction of treatment with multiple aptitudes, although not necessarily with gender.

<u>VSA and accounting education--a lab study</u>. Only one study is known to have explored a VSA-gender difference on an accounting task, although subjects were psychology students in an accounting lab setting. In an unpublished paper, Ruf,

Brown, and Crawford (1991a, 1991b) examined the moderating effect of computer training on the relationship between VSA and performance on an accounting quiz. Subjects were administered either an accounting lecture or an accounting lecture plus practice with accounting software, followed by an accounting quiz and MAP Planning Test and Mental Rotations Test. Self-reported SAT scores were used as measures of academic ability.

Results indicated that scores on the Mental Rotations Tests were related to quiz performance in both the lecture-only and computer-assisted groups. MAP Planning Test and SAT-Verbal scores, however, were not related to quiz performance in the computer-assisted group. The authors suggest that accounting software may provide visual-spatial organization of the accounting information and therefore obviate the need for the VSA factor which is measured by the MAP Planning Test.

This study also identified gender differences. For the lecture-only group, VSA and SAT scores explained more of the variance in quiz performance for males than for females. In examining the data without a gender analysis, conclusions would be (a) that more variation is accounted for by spatial relations than by spatial scanning, math ability, or verbal ability and (b) that math ability is important (Ruf, Brown, and Crawford 1991a). However, results of stepwise regression by gender showed that, for men, math ability was the only independent variable identified as important to achieving success on the quiz. For females, spatial relations (Mental Rotations Test) was most important, followed by verbal ability. In this sample, males had higher VSA and math scores than females and outperformed females on the quiz. Also, R^2 for male models was greater in simple linear (one aptitude variable at a time) and almost double in stepwise regressions the comparable R^2 for females. These results suggest the need for gender analysis in research involving VSA and academic performance.

<u>VSA--summary of research</u>. Research indicates (a) the separate existence of VSA (distinct from verbal and mathematical abilities); (b) a relationship between VSA and performance on various tasks, some obviously spatial in nature and some less so; (c) gender differences on VSA measures; and (d) gender differences on aptitude determinants for some tasks.

The next chapter discusses relevant theory and presents the five hypotheses of this dissertation. Remaining chapters present the research design and procedures, results, and discussion and conclusions.

Chapter 3 -- Theory

This chapter discusses aptitude theory and five hypotheses of the possible role of VSA in accounting education. For any given aptitude, a learning theory should specify (a) the definition of the aptitude and the measures by which it can be indicated, (b) the instructional treatments for which that aptitude is facilitative and the criterion measures by which an aptitude-treatment interaction (ATI) can be evidenced, and (c) the degree to which aptitude can be developed through training.

Aptitude constructs. Snow and Lohman (1984) discuss the aptitude theory developed by members of Stanford's Aptitude Research Project to explain the nature of cognitive aptitude for learning from instruction. Aptitude constructs describe, explain and predict the interface between the human and the task. Individuals lacking a sufficient aptitude which is requisite for a given task will perform less well on that task than will high aptitude individuals. Aptitude constructs also explain the aptitude-treatment criterion relationship. Some treatments are a function of certain aptitudes, while other treatments are not. This ATI is observable as a different relationship between the aptitude and criterion (dependent) variable for the different treatments.

Aptitude for learning from instruction describes the aptitudes which are needed for success in the learning situation. In education, describing an aptitude for learning involves describing the readiness to learn from a particular instructional treatment and the achievement criterion by which the ATI is measured. ATI is implied in all learning situations in which any aptitude, including general ability, is active.

Aptitude processes are those changes which allow learners to adapt to instructional demands. Individual differences in performance may be due to differences in (a) cognitive processes available to the individual, (b) the higher order strategic processes of organizing and monitoring the available dynamic processes, (c) the sequence of operations, and (d) the speed or efficiency of processing a sequence of operations. Individual learners may differ in speed, sequencing, availability and management of processes, and in existing knowledge and perceptions of task requirements to which the cognitive processes are applied.

Based on factor analysis of cognitive tests, aptitudes can be categorized into three primary groups which largely determine general intelligence, namely fluid intelligence, crystallized intelligence, and visual perception (Carroll 1985). Fluid intelligence, or analytic reasoning, involves demonstration of skill in induction, deduction and information processing. Crystallized intelligence, or general educational achievement, involves use of verbal knowledge and skills, as well as numerical computation (Lohman 1989). Most school learning demands both fluid and crystallized aptitude (Snow 1986). Visual perception aptitude is relevant when tasks require spatial reasoning and other visual operations; its relevance to problem-solving may be also be a function of individual differences (Snow 1986).

Whereas some tasks require the digital sequential processing of verbal-symbolic information, visual-spatial tasks generally involve the parallel processing of figural-spatial information (Snow and Lohman 1984). The theoretical model presented in Figure 1 shows that different tasks are a function of different aspects of general intelligence. Typical academic tasks, such as those involved in the SAT, are primarily a function of crystallized intelligence. To the extent that a particular task is complex and novel, fluid intelligence may also be called upon. An individual's scores on VSA tests are primarily a function of visual perception and may also be a function of fluid intelligence if the test involves tasks which are complex and novel for that individual. The model predicts that scores in academic courses may be a function of crystallized intelligence, fluid intelligence, and visual perception, depending upon the particular tasks involved in the course.

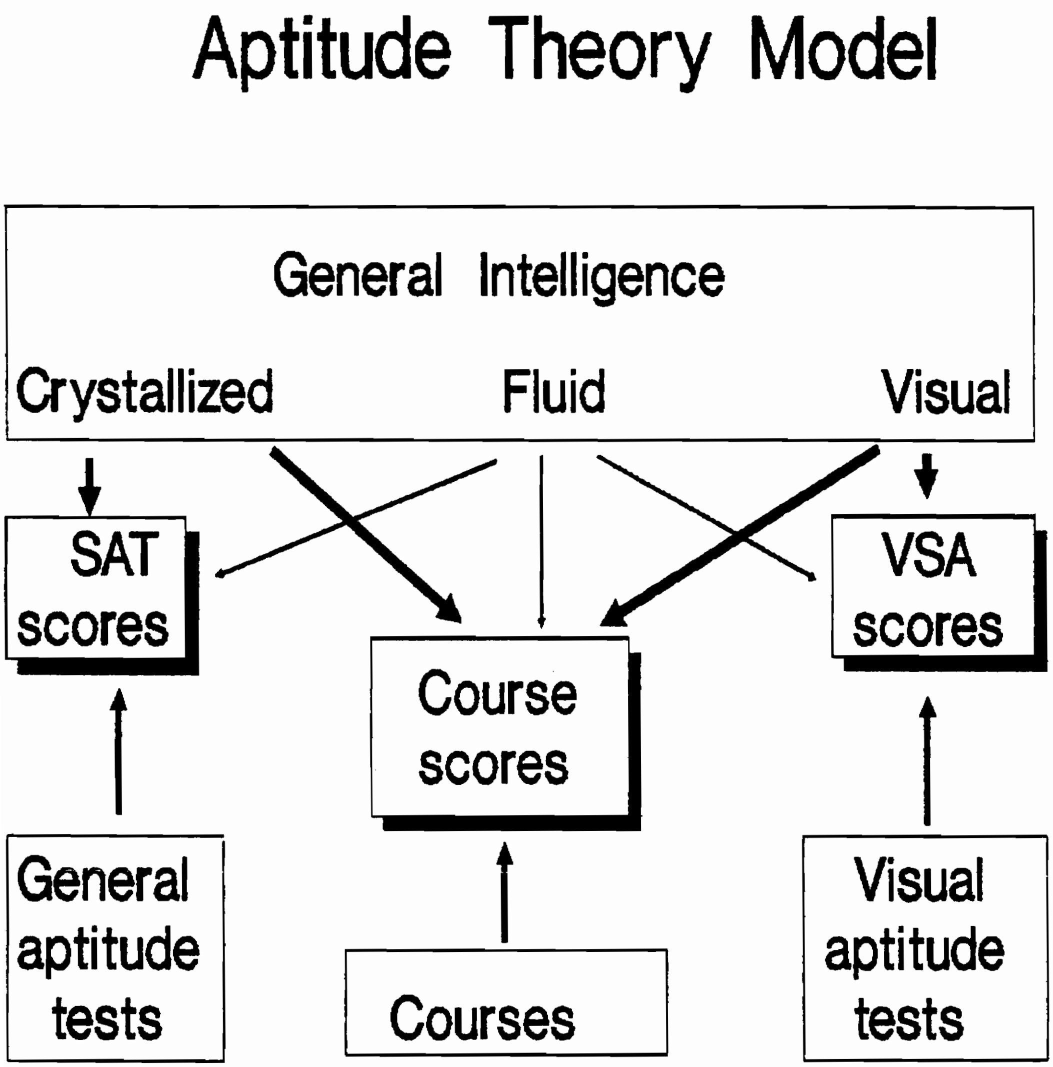


Figure 1

Aptitude Theory Model

Persons scoring similarly on a given aptitude test are expected to perform like each other on tasks for which that aptitude is important. Accordingly, the within-aptitude-group variance is hypothesized to be less than the between-aptitude-group variance on measures of task performance.

According to Pallrand and Seeber (1984), the processes associated with data analysis are related to those involved in visual-spatial thought and abstract representation. Performance in courses with data analysis burdens, such as mathematics and science, is facilitated by an ability to generate abstractions in the form of internal representations which can be altered with additional education.

VSA and accounting education. Accounting courses require processing of large amounts of data. The student is faced with an initial task of building a mental representation of a complex system designed to capture, organize and communicate financial information. Visual-spatial imagery abounds in concepts taught in this course.

Transaction information is captured using a mirror-image double-entry convention. Debits represent the left-hand entries and credits represent the right-hand entries in T-accounts, which are used in the classroom and by accountants working in business as a shorthand visual representation of the increases and decreases in account entities. The summary number of an account is termed a balance. Information flows through the system from daily lists of entries to become account histories.

The accounting equation emphasizes that assets (on the left) equal equities (on the right). Information is manipulated in generally accepted ways and presented in prescribed financial statement templates.

As the student continues to courses past first semester accounting, task requirements change. In the first accounting course, the requirement is to develop an initial understanding of the information system that is financial accounting. In the second semester course, students are introduced to cost accounting. This introduction and the subsequent Cost course seem to emphasize problem-solving. In the Intermediate course, however, the student is expected to enlarge on the initial understanding of financial accounting by learning to apply a large and complex body of rules. Research is needed to identify aspects of various courses in the accounting curriculum that may interact differentially with VSA to affect performance.

In light of the VSA-performance relationship found in other disciplines, it is expected that high-VSA students will outperform low-VSA students in coursework which requires problem-solving and data analysis. Given the nature of the mental representation task in introductory accounting, a relationship is expected between VSA and performance in that course, as evidenced by course completion rates and grades. The following hypotheses address these research questions.

Hypothesis 1. If VSA facilitates the learning of accounting, then students with high VSA should have less difficulty in completing the introductory course than should students with low VSA. This relationship should be apparent in completion rates of high and low VSA students. Specifically, there is a positive relationship between VSA and completion of the introductory accounting course.

> Students with high VSA will complete the introductory accounting course at higher rates than will students with low VSA.

This hypothesis will be tested by comparing the course completion rates of introductory accounting students categorized as high, medium or low VSA based on VSA scores, as well as by testing for VSA differences between students who drop and students who complete this course.

Hypothesis 2. If VSA facilitates the learning of accounting, VSA differences should be related to performance differences in some course grade components. Hypothesis 2 is:

> Students with high VSA will have higher performance scores in Introductory accounting than will students with low VSA.

This hypothesis will be tested by examining the linear relationships between VSA scores and grades. Models will examine both VSA alone and VSA with covariates already demonstrated to be related to accounting performance--SAT scores, college GPA, gender, the existence of prior bookkeeping coursework, and major.

Hypothesis 3. In light of research on gender differences in the VSA-performance relationship, gender differences may also be evidenced in accounting coursework. This is not to say that females and males will differ on VSA and therefore on accounting performance. Rather,

> The relationship between VSA and performance in Introductory accounting will be different for males and females.

This hypothesis will be tested by comparing the linear relationship, for each gender, for VSA and course grades. Models will examine both VSA alone and VSA with the covariates already mentioned.

Hypothesis 4. If VSA facilitates the learning of accounting, and low-VSA students encounter difficulty in introductory accounting, VSA should be related to rates of continuance to the next accounting course. Accordingly,

> Student with high VSA will continue to second semester accounting at higher rates than will students with low VSA.

This hypothesis will be tested by comparing continuance rates among students categorized as high, medium and low based on VSA test scores, as well as by comparing VSA between continuers and noncontinuers.

Hypothesis 5. VSA is expected to facilitate learning differentially depending upon the extent to which coursework requires complex mental visualization and/or emphasizes problem-solving as distinct from algorithm application.

> The relationship of VSA and performance is expected to differ among the first four accounting courses.

The relationship between VSA and performance in Intermediate accounting, which emphasizes rule application, is expected to be weaker than the relationship of VSA and Cost accounting, which emphasizes problem-solving. The direction of difference is more difficult to predict, however, for Principles I and Principles II. Principles I requires visualization of the accounting information system, while Principles II introduces cost accounting and therefore involves some problem solving. The hypothesis of course differences will be tested by comparing the linear regressions of VSA on course scores for all four courses.

The next chapter provides details of the research design for Hypotheses 1 through 5 and describes phases of data collection. Subsequent chapters report results of analyses, discussion and conclusions.

Chapter 4 -- Method

Subjects

The subjects were students enrolled in Virginia Tech's first semester Introductory accounting course in Fall 1991. Sections of Principles I are coordinated by a course manager who provides lecture notes and a common syllabus for all sections. All students take exams at common times; these exams are prepared by the course manager and are not revealed to the section instructors until immediately before exam administration. Section instructors grade exams for students in their sections following a key provided by the coordinator.

Sections are taught by the course manager and by teaching assistants, who are accounting masters and doctoral students with varying academic and business experience. Fall 1991 Principles I instructors were as follows: five male and four female teaching assistants, one female non-accounting department university staff member, and the female course manager. All but one instructor taught two sections each.

Data Collection, Phase I

On the first class session, over eight hundred students in twenty-one sections of Principles I were given a choice of participating in thirty minutes of testing. A research confederate explained that students could participate in departmentally supported research by (1) marking a decision on a consent form, (2) answering a questionnaire, and (3) taking two brief visual-spatial paper-and-pencil tests.

The consent form requested permission for the study to include (a) the student's official Virginia Tech GPA and Tech-recorded SAT scores and (b) the student's grades in Principles I and future accounting courses. Students were asked to provide their social security number and signature as evidence of their decision.

In an effort to minimize disturbances during the experimental period, students were also informed that they were welcome to fill in the questionnaire and/or the VSA tests even if they had not consented to the release of their registrar and departmental information. They were also informed that they could review the results of their VSA testing by contacting the experimenter at the end of the semester.

Demographic questionnaire. Students were then asked to answer demographic items on a questionnaire by transcribing answers onto an opscan form. Items requested historical, current, and projected student characteristics. History questions identified any previous enrollment in Principles I or a similar course, the extent of any previous accounting/bookkeeping work experience, the level of mathematics most recently passed, and whether English was the student's primary language. Current information about the student included academic rank, major field, gender, age, and primary reason for and attitude about taking Principles I. The questionnaire also asked for expected grade in Principles I as well as the student's expected number of academic hours and employment hours during that semester. Six questions permitted an answer of "other", for which instructions requested that a brief explanation be written on the face of the questionnaire in a space provided. Students were also asked to fill in their subject number on the opscan form; the subject number served to identify the section in which the student was present. A copy of the questionnaire is available in Appendix A.

VSA assessment. VSA was assessed by the MAP Planning Test (MAP) (Ekstrom, French, and Harman 1976), a measure of spatial scanning, and the Mental Rotations Test (MRT) (Vandenburg and Kuse 1978), a measure of spatial relations. MAP and MRT each consist of a brief instruction and practice followed by two 3-minute pencil-and-paper tasks. MAP is available as part of the ETS Kit of Factor-Referenced Cognitive Tests. It requires subjects to trace a pathway past checkpoints on a grid. MRT requires the matching of pictures of three-dimensional block objects, rotated in at least one plane, to a target object. MAP and MRT scores are continuous measures ranging from 0 to 40. See Appendix B for sample pages and answer keys for each VSA instrument. The authors report test-retest reliability of .80 for the MAP (Ekstrom et al. 1976) and .83 for the MRT and an internal consistency reliability coefficient of .88 for the MRT (Vandenberg and Kuse 1978).

Processing of questionnaires. An electronic data file was generated from opscan responses, stripped of records of nonconsenters, and transmitted to the Registrar's office where SAT scores and GPA were added. This file was returned electronically to the experimenter's computer account, and the consent forms were permanently filed with the experimenter.

Processing of VSA responses. Twelve data items were recorded for the hand-scored MAP and MRT. Items recorded for each test were the subtotals right and wrong for each of two parts as well as the total right and total wrong. Every record was then checked against the hand-scored sheets.

Data Collection, Phase II

Immediately after the drop date, post-drop class rosters were compared with pre-drop rosters. Subjects who dropped from Principles I were identified, and their records were marked accordingly for use in analysis of Hypothesis 1.

Data Collection, Phase III

Toward the end of the semester, data were collected on the instructors, including MAP and MRT assessment of VSA and a questionnaire on educational background, accounting and teaching experience, age, intended career (academic or nonacademic). Instructors were also asked to choose among seven statements concerning their attitude toward teaching sections of a coordinated course. Preference scale endpoints ranged from having sole responsibility for a course to teaching within a coordinated course.

Data Collection, Phase IV

After the semester, Principles I performance measures were obtained from section instructors for all consenting subjects. The measures were scores on (1) homework and quizzes, (2) a computerized practice set, and (3) three exams and a final exam. The items were entered into a computer file, and each record checked against the photocopies of grades provided by instructors. The four exams were reviewed by the experimenter to identify differences in content and question type.

Data Collection, Phase V

Longitudinal data collection addressed enrollment and performance in accounting courses subsequent to Principles I. Principles II rosters were examined to identify subjects who continued to Principles II. At the end of the semester, photocopies or electronic grade files of grades were secured from section instructors.

The process was repeated in Fall 1992 to determine whether completers of Principles II enrolled in Cost or Intermediate Accounting and what grades the enrollees made in the first two exams in these courses. The first exam in each course is considered to be at least a partial review of previous accounting courses. The second exam is the first measure thought to test new material related specific to that course.

Analysis

All data were analyzed using SAS, release 6.06. Due to unequal cell sizes, PROC GLM was used for MANOVA, MANCOVA, and regression analyses.

Chapter 5 -- Results

Chapter organization

Research results are reported in eight major sections. First, sample mortality is described. Second, descriptive statistics are provided for the aptitude and demographic variables. Third, results are reported for tests of Hypothesis 1, which predicts a relationship between VSA and rates of completing Principles I. Fourth, demographic variables are evaluated as possible covariates for tests of remaining hypotheses. The last four sections report results from tests of Hypotheses 2 through 5. Each chapter section is further divided into labelled subsections.

Sample Description

Of the students present in first-day classes who were given the option to participate in the research, 708 (85%) agreed to take the VSA tests, answer the demographic questionnaire and provide written consent for experimenter access to Registrar records of student GPA and SAT scores and to departmental records of grades in Accounting courses. Subsequent analysis resulted in exclusion of 132 subjects' data for several reasons. Some students were auditing or not enrolled in the class in which they were physically present. Records are also excluded if they have zero (or missing) values for aptitude or demographic variables.

Table 1 reports the mortality figures for the sample, explaining the incremental sample reduction to the 576 records in the initial sample used in analysis of Hypothesis 1 and subsequently to the 513 in the primary sample used for analysis of Hypotheses 2 and 3. Table 1 also reports the incremental reduction of sample size resulting from students not enrolling in subsequent courses in the accounting series. Students who completed Principles II are in one of four categories the following fall: enrolled in both Cost and Intermediate accounting, enrolled only in Cost, enrolled only in Intermediate, enrolled in neither.

Table 1
Mortality figures for sample reduction

History of Primary Sample	#	percentage of 830	708	576
Had option to participate	830			
less nonparticipating	122	15%		
Consented to participate	708	85%		
less missing VSA scores, zero/missing GPA or SAT scores, gender not reported, and/or audit only or not enrolled	132		19%	
Records used in analysis of Hypothesis 1	576		81%	
less dropped or grades unavailable	63			11%
Records used in analysis of Hypotheses 2, 3	513			89%
less those not passing Principles I	41			
Records used in analysis of Hypothesis 4	472			
less those not continuing to Prin. II	130			
Records used in analysis of Hypothesis 5	342			

Of eligible subjects, enrollment after Principles II was: 169 in neither course, 94 in Intermediate, 75 enrolled in Cost, 71 enrolled in Cost and Intermediate.[3]

[3]Total initial enrollment in Intermediate sections was 234 and in Cost was 242, although a number of these students dropped throughout the semester.

Demographics

Tables 2 and 3 report the frequencies for categorical questionnaire variables for the initial sample (N = 576) used for Hypothesis 1. In the following discussion, and throughout the results section, variables are in capital letters.

The most common profile for Principles I students was that of an 18-19 year old sophomore, native English speaker, with no Principles I previous enrollment and no accounting work experience. Most subjects report calculus as their highest math class passed, and most anticipate no employment hours during the semester. Subjects' majors are predominantly accounting, general business or other business.

Whereas 61% of subjects claim no prior bookkeeping/ accounting course of any kind, 39% report one or more such courses in high school or another university. For most subjects, the primary reason for enrollment in Principles I is that it is a requirement for their majors, and most anticipate a grade of A or B. Subjects report feelings about taking Principles I as excited (10%), interested (45%), mildly curious (10%), neutral (21%), resigned (2%), and worried (12%). The initial sample is 47% females and 53% males.

Table 2
Demographic variables--frequencies and percentages
RANK, MAJOR, GENDER, AGE, PRINPREV, PRIOR, WORK
N=569 (initial sample used for drop analysis)

Variable	Values	Freq.	%
RANK	Freshman	41	7.1
	Sophomore	498	86.5
	Juniors	30	5.2
	Seniors or others	7	1.2
MAJOR	Accounting	140	24.3
	Agricultural Economics	12	2.1
	Business	123	21.4
	Economics	11	1.9
	Finance	66	11.5
	Management	44	7.6
	Management Science	30	5.2
	Marketing	59	10.2
	None or other	91	15.8
GENDER	Females	269	47.0
	Males	307	53.0
AGE	18	124	21.5
	19	382	66.3
	20	53	9.2
	21-24	17	3.0
PRINPREV	Not previously enrolled in course	557	96.7
	Previously enrolled, but dropped	2	.3
	Previously enrolled, made a C	12	2.1
	Previously enrolled, made a D	2	.3
	Previously enrolled, made an F	5	.9
PRIOR	No prior bookkeeping course	349	60.7
	One prior course in high school	145	25.2
	Two or more prior in high school	54	9.4
	One prior course in college	22	3.8
	Two or more prior in college	5	.9
WORK	No accounting work experience	497	86.3
	Less than 6 months accounting work	63	10.9
	More than 6 months accounting work	16	2.8

Table 3
Demographic variables--frequencies and percentages
MATHMAX, REASON, GRADE, FEELING, HOURS, SEMWORK
N=569 (initial sample used for drop analysis)

Variable	Values	Freq.	%
MATHMAX	Algebra	3	.5
	Trigonometry	6	1.0
	Calculus	564	98.5
REASON	Required for Major	438	76.3
	Curious	34	5.9
	Prerequisite for other courses	13	2.3
	Get into School of Business	78	13.6
	Other	11	1.9
GRADE	Expect an A in Principles I	127	22.2
	Expect a B in Principles I	364	63.5
	Expect a C in Principles I	79	13.8
	Expect a D in Principles I	2	.3
	Taking course Pass/Fail	0	
	Auditing course	1	.2
FEELING	Excited about Principles I	54	9.4
	Interested	260	45.4
	Mildly curious	56	9.8
	Neutral	118	20.6
	Resigned	13	2.3
	Worried	70	12.2
	Resentful	2	.3
HOURS	Enrolled for fewer than 12	17	3.0
	Enrolled for 12-15	337	56.4
	Enrolled for 16-18	221	38.6
	Enrolled for 19 or more	12	2.1
SEMWORK	Not employed this semester	399	69.8
	Employed fewer than 10 hours/week	52	9.1
	Employed 10-19 hours/week	90	15.7
	Employed 20-29 hours/week	26	4.5
	Employed 30-39 hours/week	2	.3
	Employed 40 hours or more/week	3	.5

Descriptive statistics for aptitude variables

VSA measures used in analysis include the total number right and total number wrong on the MAP Planning test, MAP-RIGHT and MAP-WRONG, and on the Mental Rotations Test, MRT-RIGHT and MRT-WRONG. VSA variables also include the total right as a percentage of items attempted on the MAP Planning Test, MAP-%, and on the Mental Rotations Test, MRT-%. These percentage variables serve as an accuracy measure. Their use minimizes the effects of differing speed and cautiousness of subjects. Correction-for-guessing formulas, which assume one correct answer for each item, were not used due to the format of the MRT which requires two answers for each item. Other aptitude measures are the two Scholastic Aptitude Test scores, SAT-MATH and SAT-VERBAL, and grade point average, GPA. Table 4 reports the means, standard deviations, and ranges for the VSA measures, GPA, and SAT scores for the primary sample.

Table 4
Aptitude means, standard deviations, and ranges for the primary sample

N = 513

variable	mean	(s.d.)	minimum	maximum
MAP-RIGHT	24.24	(6.58)	0.00	40.00
MAP-WRONG	1.82	(2.69)	0.00	23.00
MAP-%	92%	(.12)	0.00	1.00
MRT-RIGHT	22.58	(8.12)	3.00	40.00
MRT-WRONG	4.00	(3.67)	0.00	26.00
MRT-%	84%	(.14)	.23	1.00
SAT-MATH	572.11	(70.48)	390.00	800.00
SAT-VERBAL	481.27	(68.67)	310.00	680.00
GPA	2.50	(.57)	.50	4.00

Tests of Hypothesis 1

Hypothesis 1 predicts a relationship between VSA and completing Principles I. Tests of this hypothesis are by discriminant analysis, *t*-test, and Chi-square.

Discriminant analysis. In multivariate discriminant analysis, multiple aptitude variables are used to predict an individual's identity as either a dropper or completer. Data for subjects were sorted by social security number, then divided into either analysis or test (holdout) samples depending on comparison of their social security number with the criterion number which allocated the lowest 60% into the analysis group and the highest 40% into the test group.

For each model, a discriminant function is developed with the analysis sample and then applied to the holdout sample. The resulting classification of holdout subjects as droppers or completers is then compared to the actual identity of these subjects as droppers and completers. The ratio of correct classification (hit ratio) is calculated as the number correctly classified divided by the number classified.

This classification ratio is compared to the ratio achievable by chance given the ratio of actual droppers to completers in the analysis sample. First, the choice of the appropriate chance criterion must be made. If there were equal numbers of droppers and completers, then a maximum chance criterion would be appropriate (Hair, Anderson and

Tatham 1987).

A maximum chance criterion (C_{Max}) is the percent which would be correctly classified if all records were classified into the larger group (if groups are unequal) or consistently into either group (if groups are equal). Groups for the current purpose are droppers and completers. Since the numbers of droppers and completers are not equal, however, the chance ratio should instead be a proportional criterion (C_{Pro}), calculated as the sum of the squared proportion of droppers and the squared proportion of completers. The C_{Pro} for the analysis sample is .8548.

Table 5 reports the classification matrices and rates for droppers and completers for a series of discriminant function models. Models predict drop-complete classification based (1) solely on VSA components, (2) solely on VSA percentages, (3) solely on SAT scores, (4) on SAT scores and GPA, and (5 and 6) on SAT scores, GPA and VSA scores (either as components or percentages). For all models, when the discriminant function developed with the analysis sample is applied to data in the holdout sample, over 90% of classifications are correct, which is a result greater than C_{Pro} = .8548. However, this rate is achieved because all of the functions developed by the discriminant analysis are classifying all subjects as completers, thereby attaining a C_{Max} hit rate. For discriminant analysis to be a valid statistical tool,

Table 5
Results of discriminant analyses of aptitudes to differentiate Principles I droppers and completers

Classification Matrices for Holdout Sample (N = 235)
Proportional Chance Ratio = .8595 [4]

Variables in Model	Identity	Classification Comp.	Classification Drop.	Hit[5] Ratio
MAP-RIGHT, MAP-WRONG,	Comp.	219	0	
MRT-RIGHT, MRT-WRONG	Drop.	16	0	.9319
	Comp.	219	0	
MAP-%, MRT-%	Drop.	16	0	.9319
	Comp.	219	0	
SAT-MATH, SAT-VERBAL	Drop.	16	0	.9319
SAT-MATH,	Comp.	219	0	
SAT-VERBAL, GPA	Drop.	16	0	.9319
All aptitudes	Comp.	219	0	
(VSA, SATs, GPA)	Drop.	16	0	.9319

information provided by variable levels must be used to categorize subjects into groups (Hair, Anderson and Tatham 1987). The discriminant functions described in the table are not valid discriminators between the groups of interest because they assign all subjects to only one group. Discriminant analysis is not usable to test Hypothesis 1.

[4] The proportional chance ratio (C_{Pro}) is calculated as the sum of the square of the proportion of completers plus the square of the proportion of droppers.

[5]Hit ratio is calculated as the number correctly classified divided by the number classified. This is compared to the proportional chance ratio.

T-test results, however, indicate that droppers do differ from completers on the basis of the aptitudes examined in this study. Table 6 reports aptitude means for each group and *t*-test results. Those subjects who dropped Principles I have lower average scores on MAP-RIGHT, MRT-RIGHT, SAT-MATH, SAT-VERBAL, and GPA. This finding supports Hypothesis 1.

Table 6
Aptitude means, standard deviations and
t-test values for Principles I droppers and completers

Variable	Completers N=533 Means (s.d.)	Droppers N= 43 Means (s.d.)	*t*-test values
MAP-RIGHT	24.08 (6.65)	21.86 (6.04)	2.12**
MAP-WRONG	1.84 (2.69)	2.05 (3.35)	- .39
MAP-%	.92 (.12)	.92 (.12)	.32
MRT-RIGHT	22.54 (8.08)	19.81 (7.58)	2.14**
MRT-WRONG	4.02 (3.68)	3.93 (3.42)	.16
MRT-%	.84 (.14)	.83 (.12)	.50
SAT-MATH	569.91 (71.76)	544.19 (75.73)	2.25**
SAT-VERBAL	481.29 (69.42)	453.72 (66.62)	2.51**
GPA	2.49 (.57)	2.07 (.52)	4.63****

[**** p = .0001, *** p = .001, ** p < .05, * p < .10]

This is a different conclusion, of course, from asserting that lower aptitude students dropped more often than higher aptitude students. To test this assertion, Chi-square tests determined if drop frequency is independent of three VSA levels. Subjects were ranked by MAP-RIGHT and separately by MRT-RIGHT. The highest, middle and lowest thirds of each variable form analysis groups. Cut-off scores for each group were chosen from among ranked values such that approximately equal numbers of subjects would fall within each of three categories for each variable. As a result, MAP-RIGHT scores between 22 and 26 inclusive comprised middle MAP-RIGHT, and scores above and below these scores comprised the high and low MAP-RIGHT categories For MRT-RIGHT, scores between 19 and 25 inclusive comprised the middle group, and scores above and below these scores comprised the high and low MRT-RIGHT categories. Subjects were not constrained to be in comparable groups for the two variables.

Table 7 reports Chi-square values which indicate that drop rates are not independent of three levels of MAP-RIGHT and not independent of three levels of MRT-RIGHT. These relationships are evident for the high and low groups. Using the proportion of droppers to completers in the overall sample, we would expect 20 or 21 droppers in each category. Actual frequencies, however, are 27 (low MAP-RIGHT), 24 (medium MAP-RIGHT), and 11 (high MAP-RIGHT).

Table 7
Summary of results of
Chi-square tests of independence
between drop rates and VSA levels[6]

Cell values:
actual frequency
[expected frequency]

	Droppers	Completers	Chi-square (p-value)
MAP-RIGHT			
low	27 [20.487]	163 [169.51]	
medium	24 [20.918]	170 [173.08]	
high	11 [20.595]	180 [170.41]	7.84 ($p < .02$)
MRT-RIGHT			
low	29 [21.547]	170 [177.54]	
medium	20 [20.056]	166 [165.94]	
high	13 [20.487]	177 [169.51]	6.04 ($p < .05$)

That is, more low-MAP subjects drop than expected, and fewer high-MAP subjects drop than expected, given the proportion of droppers to completers. Actual frequencies for MRT-RIGHT are 29 (low), 20 (medium) and 13 (high). More low-MRT subjects drop than expected, and fewer high-MRT subjects drop than

[6]Chi-square results for MAP-WRONG and MRT-WRONG were not significant.

expected.

Tests of Hypothesis 1--summary. Chi-square results indicate a relationship between VSA levels and completion of Principles I. Results of t-tests indicate that completers have higher VSA than do droppers of Principles I. These findings support the hypothesized relationship between VSA and completing behavior.

Demographic-performance analysis

Before testing Hypotheses 2, 3 and 5, demographic variables are examined for possible use in the tests of hypotheses. First, gender differences in performance are explored by t-tests. Second, other demographic differences in performance are identified using Tukey analysis. Third, for each demographic variable, MANOVA examines group differences in the vector of Principles I grades.

Performance measures. Dependent variables are the six primary grade scores which determine the Principles I semester grade: four exams with maximum values of 100, 100, 100 and 150 points (EXAMS 1-4), a homework/quiz (HW/QUIZ) score of 40 maximum points, and a computerized practice set (COMPUTER) with a maximum value of 60.[7]

---

[7]The practice set used in Principles I and II is It's Your Business (Business Publications, Inc., 1987) by Donald V. Saftner and Rosalind Cranor, designed for use with the software BGL (Basic General Ledger).

Performance t-tests by gender. For the first demographic variable, GENDER, Table 8 reports performance means and standard deviations separately for females and males, and provides p-values for t-tests for gender differences. Although there are no performance differences on EXAM 1, females score higher on EXAM 2, and lower on EXAM 3 and EXAM 4. Females also score higher on HW/QUIZ and COMPUTER. There is, however, no gender difference in total score and therefore the final course grade.

Table 8
Female and male Principles I performance score means standard deviations and results of *t*-tests across genders

N=513 (239 females, 274 males)

Means (s.d.)

	Female	Male	t-test value
EXAM 1	79.93 (12.61)	80.25 (11.08)	- .31
EXAM 2	74.87 (14.84)	71.68 (14.80)	2.44**
EXAM 3	74.13 (14.93)	77.02 (13.21)	-2.31**
EXAM 4	92.10 (26.32)	97.10 (23.76)	-2.26**
HW/QUIZ	36.41 (6.04)	33.68 (8.07)	4.37****
COMPUTER	55.89 (6.53)	52.93 (9.73)	4.08****
TOTAL	413.32 (66.45)	412.66 (63.67)	.11

Tukey analysis of performance by other demographic variables. For nine demographic variables, Tukey analysis of performance mean differences across demographic categories shows neither univariate differences on any performance component nor multivariate differences. These variables for which no reliable differences in performance are found include academic rank, English as a primary or second language, previous enrollment in Principles I, previous accounting work experience, highest math course passed, reason for taking Principles I, number of academic hours and number of employment hours during the Principles I semester. A trivial performance difference which was found for expected grade identified that on EXAM 1 the one student expecting a D in the course got a lower score than all other students. Due to the small N expecting a D, this variable is classified with those for which no performance differences were found.

In Tables 9 and 10, Tukey lines are presented showing three other demographic variables and their values for which differences in performance are found in Tukey analysis. Table 9 Panel A reports differences found for FEELING, for which reliable differences are found only on HW/QUIZ and COMPUTER. For HW/QUIZ, worried students (N = 57) scored higher than resigned students (N = 7). Otherwise, subjects scoring at the other five points of the scale do not perform differently from each other on HW/QUIZ. For COMPUTER, excited (N = 52), worried (N = 57), and interested (N = 236) subjects scored higher than resigned (N = 10) and resentful (N = 2) subjects. In both HW/QUIZ and COMPUTER, however, the majority of subjects fell into FEELING categories which did not perform differently from each other.

Table 9 Panel B reports on Tukey differences found on the presence or absence of a PRIOR bookkeeping course. Those students who report taking one or more prior bookkeeping or accounting courses (PRIOR) in high school or at another college (N = 214) outperform nontakers (N = 299) on EXAM 1, EXAM 2 and EXAM 4, but not on EXAM 3 or HW/QUIZ or COMPUTER.

Table 9
Tukey comparison of Principles I performance scores by FEELING and PRIOR bookkeeping course

PANEL A:

FEELING (feelings about taking Principles I: excited, interested, mildly curious, neutral, resigned, worried, resentful) [Mean differences on HW/QUIZ and COMPUTER]

Tukey grouping		HW/QUIZ Mean	N	Feeling
	A	36.544	57	worried
B	A	35.596	52	excited
B	A	35.398	236	interested
B	A	35.059	55	mildly curious
B	A	33.606	104	neutral
B	A	28.500	2	resentful
B		26.500	7	resigned

Tukey grouping		COMPUTER Mean	N	Feeling
	A	56.019	52	excited
	A	55.211	57	worried
	A	55.127	236	interested
B	A	54.588	51	mildly curious
B	A	52.135	104	neutral
B		44.000	10	resigned
B		44.000	2	resentful

PANEL B:

PRIOR (other prior bookkeeping/accounting course: no, yes) [Mean of subjects answering yes higher on EXAMS 1,2,4.]

Tukey group		EXAM 1 Mean	N	Prior
	A	83.26	214	yes
B		77.84	299	no

Tukey group		EXAM 2 Mean	N	Prior
	A	75.84	214	yes
B		71.25	299	no

Tukey group		EXAM 4 Mean	N	Prior
	A	97.49	214	yes
B		92.82	299	no

Table 10 reports Tukey differences found on academic MAJOR. Accounting majors (N = 133) reliably outperform other business majors (N = 304) but not nonbusiness majors (N = 76) on all exams (EXAMS 1-4), whereas accounting majors outperform nonbusiness majors but not other business majors on COMPUTER. The HW/QUIZ mean of accounting majors is not significantly higher than that of other business majors nor is the mean of the business majors significantly higher than that of nonbusiness majors.

Table 10
Tukey comparison of Principles I performance scores by MAJOR

N = 513

MAJOR (accounting, other business, nonbusiness)

Tukey group		EXAM 1 Mean	N	Major
	A	83.86	133	accounting
B	A	81.51	76	nonbusiness
B		78.11	304	other business

Tukey group		EXAM 2 Mean	N	Major
	A	78.53	133	accounting
B		72.32	76	nonbusiness
B		71.03	304	other business

Tukey group		EXAM 3 Mean	N	Major
	A	79.20	133	accounting
B	A	75.65	76	nonbusiness
B		74.14	304	other business

Tukey group		EXAM 4 Mean	N	Major
	A	102.41	133	accounting
B	A	96.86	76	nonbusiness
B		91.91	304	other business

Tukey group		COMPUTER Mean	N	Major
	A	56.20	133	accounting
B	A	53.69	304	other business
B		53.49	76	nonbusiness

<u>MANOVA of performance scores on demographic variables</u>. Tables 11 and 12 report, for the whole sample and for gender sample subsets, the R^2 and p-values for those demographic models for which either a univariate or multivariate test is significant at conventional levels. Academic RANK is related to the grade vector and to all grade components except EXAM 2 and EXAM 4, with R^2 ranging from around .01 to almost .04. PRINPREV is related to the grade vector, but univariately significant only on HW/QUIZ, with R^2 around .03. MATHMAX is related to all EXAMS, with R^2 ranging from around .02 to almost .04. SEMWORK is related to the grade vector and to EXAM 1 and HW/QUIZ, with R^2 ranging from .02 to .03.

Variables related to the grade vector at $p < .0001$ include MAJOR, PRIOR, GRADE, and FEELING. MAJOR is positively related to all EXAMS and COMPUTER. R^2 ranges from around .01 to almost .05. PRIOR is related to EXAMS 1, 2 , and 4. R^2 ranges from less than .01 to more than .05. Expected GRADE, which is positively related to HW/QUIZ and COMPUTER, is excluded as a possible demographic candidate due to the small N referred to above.

Table 11
Results of MANOVA of Principles I performance scores by RANK, MAJOR, PRINPREV, PRIOR
[4 of the 8 demographic models significant at $p < .10$]
N = 513 (239 females, 274 males)

Model	All p <	All [R^2]	Female p <	Female [R^2]	Male p <	Male [R^2]
RANK:						
multivariate	.0115		.3754		.0101	
EXAM 1 = RANK	.0419	[.0193]	.3833	[.0176]	.1051	[.0224]
EXAM 2 = RANK	.3078	[.0094]	.2346	[.0234]	.2434	[.0153]
EXAM 3 = RANK	.0426	[.0192]	.1768	[.0265]	.0412	[.0300]
EXAM 4 = RANK	.1154	[.0145]	.2762	[.0215]	.0938	[.0234]
HW/QUIZ= RANK	.0029	[.0312]	.0998	[.0326]	.0124	[.0394]
COMPUTER=RANK	.0006	[.0381]	.5677	[.0124]	.0002	[.0710]
MAJOR:						
multivariate	.0001		.0007		.0144	
EXAM 1= MAJOR	.0001	[.0454]	.0021	[.0508]	.0013	[.0477]
EXAM 2= MAJOR	.0001	[.0465]	.0003	[.0654]	.0147	[.0307]
EXAM 3= MAJOR	.0024	[.0233]	.0027	[.0488]	.0141	[.0310]
EXAM 4= MAJOR	.0001	[.0392]	.0001	[.1074]	.1584	[.0135]
HW/QUIZ=MAJOR	.1095	[.0086]	.9928	[.0000]	.0669	[.0198]
COMPUTER=MAJOR	.0118	[.0173]	.2084	[.0132]	.0706	[.0194]
PRINPREV:						
multivariate	.0050		.7572		.0011	
EXAM 1 = PRIN	.1015	[.0151]	.7485	[.0052]	.0130	[.0391]
EXAM 2 = PRIN	.4053	[.0078]	.7214	[.0056]	.2119	[.0165]
EXAM 3 = PRIN	.7449	[.0038]	.7941	[.0044]	.4696	[.0093]
EXAM 4 = PRIN	.6732	[.0046]	.6660	[.0066]	.2362	[.0156]
HW/QUIZ= PRIN	.0024	[.0319]	.1326	[.0235]	.0021	[.0528]
COMPUTER=PRIN	.1707	[.0125]	.4963	[.0101]	.1861	[.0176]
PRIOR:						
multivariate	.0001		.0016		.0001	
EXAM 1 = PRIOR	.0001	[.0513]	.0009	[.0459]	.0001	[.0586]
EXAM 2 = PRIOR	.0005	[.0232]	.0064	[.0310]	.0447	[.0147]
EXAM 3 = PRIOR	.3283	[.0019]	.4919	[.0020]	.3689	[.0030]
EXAM 4 = PRIOR	.0376	[.0084]	.0080	[.0293]	.6622	[.0007]
HW/QUIZ= PRIOR	.9804	[.0000]	.2835	[.0049]	.6245	[.0009]
COMPUTER=PRIOR	.3570	[.0017]	.7221	[.0005]	.2856	[.0042]

Table 12
Results of MANOVA of Principles I performance scores
by MATHMAX, GRADE, FEELING, SEMWORK
[4 of the 8 demographic models significant at p < .10]
N = 513 (239 females, 274 males)

Model	All p <	All [R^2]	Female p <	Female [R^2]	Male p <	Male [R^2]
MATHMAX:						
multivariate	.0678		.0001		.0883	
EXAM 1 = MATH	.0042	[.0332]	.4172	[.0211]	.0048	[.0539]
EXAM 2 = MATH	.0266	[.0247]	.2308	[.0290]	.1212	[.0266]
EXAM 3 = MATH	.0960	[.0183]	.8083	[.0097]	.0489	[.0347]
EXAM 4 = MATH	.0016	[.0377]	.0776	[.0416]	.0120	[.0465]
HW/QUIZ= MATH	.4395	[.0094]	.6423	[.0144]	.4917	[.0126]
COMPUTER=MATH	.1880	[.0146]	.0001	[.1193]	.4551	[.0134]
GRADE:						
multivariate	.0001		.0002		.0001	
EXAM 1 = GRADE	.0001	[.0979]	.0001	[.0920]	.0001	[.1085]
EXAM 2 = GRADE	.0001	[.0791]	.0001	[.1114]	.0001	[.0807]
EXAM 3 = GRADE	.0001	[.0848]	.0001	[.0770]	.0001	[.0835]
EXAM 4 = GRADE	.0001	[.1138]	.0001	[.1079]	.0001	[.1072]
HW/QUIZ= GRADE	.8896	[.0012]	.9830	[.0001]	.8711	[.0026]
COMPUTER=GRADE	.2010	[.0091]	.4980	[.0060]	.1913	[.0174]
FEELING:						
multivariate	.0001		.0068		.1818	
EXAM 1 = FEEL	.0001	[.0545]	.0004	[.0928]	.2498	[.0287]
EXAM 2 = FEEL	.0017	[.0409]	.0146	[.0589]	.1673	[.0333]
EXAM 3 = FEEL	.0412	[.0255]	.1806	[.0320]	.2404	[.0292]
EXAM 4 = FEEL	.0002	[.0500]	.0120	[.0608]	.0261	[.0519]
HW/QUIZ= FEEL	.0009	[.0438]	.0265	[.0529]	.1014	[.0387]
COMPUTER=FEEL	.0001	[.0573]	.0088	[.0638]	.0304	[.0505]
SEMWORK:						
multivariate	.0044		.0341		.0825	
EXAM 1 = SEM	.0325	[.0237]	.4366	[.0205]	.0197	[.0424]
EXAM 2 = SEM	.8623	[.0037]	.7929	[.0102]	.5365	[.0115]
EXAM 3 = SEM	.0563	[.0210]	.3850	[.0223]	.1735	[.0233]
EXAM 4 = SEM	.5210	[.0082]	.7126	[.0124]	.2516	[.0197]
HW/QUIZ= SEM	.0099	[.0293]	.0029	[.0744]	.1051	[.0280]
COMPUTER=SEM	.4111	[.0099]	.1021	[.0385]	.2784	[.0187]

Demographic variables chosen for further analysis. GENDER, PRIOR, MAJOR and FEELING are the only variables for which consistent results are found for gender t-tests and MANOVA and Tukey analysis of other demographic variables. FEELING is excluded from consideration as a demographic variable due to concerns about its stability over the period in which performance measures are taken. It represents a student's feeling at one point in time, not over the course of the semester. It is not reasonable to suppose that a student's initial feeling concerning a course will be the same over four months experience with the course. Accordingly, GENDER, PRIOR and MAJOR will be used as demographic variables (with SAT-MATH, SAT-VERBAL and GPA as continuous covariates) in direct testing of Hypothesis 2. Only PRIOR and MAJOR will be used as demographic variables in testing Hypothesis 3, since separate models are developed by GENDER to test gender differences in the aptitude-performance relationship.

Tests of Hypothesis 2

Hypothesis 2 posits a relationship between VSA and performance in Principles I. This hypothesis is explored in two primary approaches. Multivariate analysis of variance (MANOVA) and covariance (MANCOVA) treat MAP-RIGHT and MRT-RIGHT as categorical after the sorted values have been divided into three approximately equal groups for each variable.

Second, regressions are examined which either include MAP-RIGHT, MAP-WRONG, MRT-RIGHT and MRT-WRONG or include MAP-% and MRT-%. One set of regressions examines the significance of VSA measures alone in models referred to as partial models. The second set of regressions (termed full models) and MANCOVAs include the demographic variables selected above and SAT-MATH, SAT-VERBAL, and GPA. Before these analyses are reported, however, correlations among aptitude variables and between aptitude variables and performance are discussed.

<u>Correlation among aptitude variables</u>. One of the assumptions of MANCOVA and regression is independence of independent variables. If the independent variables are highly correlated, regression coefficients may be incorrectly estimated and have the wrong signs (Hair, Anderson and Tatham 1987). It is also difficult to separate the explanatory contribution of each variable (Ott 1988).

Table 13 reports correlations for aptitude measures for the primary sample. Correlations calculated for females only and for males are not reported in a separate table but are noted in Table 13 as follows; those correlations which are also significant for both females and males are marked as follows: f for female significance at $p < .10$, F for female significance at $p < .05$, m for male significance at $p < .10$, and M for male significance at $p < .05$.

None of the aptitude variables are highly correlated

Table 13
Correlations among aptitudes,[8]
primary sample

N = 513

	MAP-RIGHT	MAP-WRONG	MAP-%
MAP-RIGHT			
MAP-WRONG	-.4662****F M		
MRT-RIGHT	.3240****F M	-.0953** F	
MRT-WRONG	-.0928** M	.1351** M	
SAT-MATH	.2772****F M	-.0903** F M	.1171** M
SAT-VERBAL	.0487	-.0932** M	.0939** M
GPA	.0087	-.0362 M	.0249
MRT-%			.1710****F M

	MRT-RIGHT	MRT-WRONG	MRT-%
MAP-RIGHT	.3240****F M	-.0928**	
MAP-WRONG	-.0953** F	.1351** M	
MRT-RIGHT		-.2492****F M	
MRT-WRONG	-.2492****F M		
SAT-MATH	.2232****F M	-.2156****F M	.2547****F M
SAT-VERBAL	.1228** M	-.0786* M	.1174** M
GPA	-.0146	-.1289** M	.1319** M
MAP-%	.1297** F	-.1467*** M	.1710****F M

	SAT-MATH	SAT-VERBAL	GPA
SAT-MATH		.3778****F M	.2341****F M
SAT-VERBAL			.2519****F M

[**** p = .0001, *** p = .001, ** p < .05, * p < .10]

with each other. Only three correlations exceed .30. MAP-RIGHT is correlated with MRT-RIGHT at .3240; MAP-RIGHT is correlated with MAP-WRONG at -.4662; and SAT-MATH is correlated with SAT-VERBAL at .3778.

[8]Females: F = sig. (p < .05) f = sig. (p < .10)
Males: M = sig. (p < .05) m = sig. (p < .10)

Aptitude-performance correlations. Table 14 reports that correlations between VSA measures and performance scores are low. There is only one correlation between MAP and a performance score significant at $p < .05$: MAP-RIGHT and EXAM 1, correlated at .0909. MRT-RIGHT is significantly correlated with EXAM 3 (correlation .0953) and EXAM 4 (correlation .0990). MRT-WRONG is correlated at p-values ranging from .0001 to .05 for all performance scores except COMPUTER; correlations are -.1460 (EXAM 1), -.1135 (EXAM 2), -.1706 (EXAM 3), -.1727 (EXAM 4), HW/QUIZ .0917 (HW/QUIZ), and -.1387 (Principles I TOTAL1). MRT-% has comparable but opposite sign correlations with these same grades. From these correlations, it might be expected in continuous analysis that a relationship with MRT and performance will more likely to be demonstrated than a MAP-grade relationship, and a likelihood that MRT-WRONG will be more explanatory than MRT-RIGHT.

SAT-VERBAL is positively correlated with all EXAMS and TOTAL (correlations ranging from .1441 to .2223) and negatively correlated with HW/QUIZ (correlation -.1348) and COMPUTER (correlation -.0964). SAT-MATH has the same pattern, with EXAM correlations ranging from .2056 to .3641. GPA is correlated at $p < .0001$ with all grades, correlations ranging from -.4387 (EXAM 2) to .5224 (TOTAL).

The negative relationship between most aptitude variables and HW/QUIZ and COMPUTER is consistent with the

Table 14
Correlations among aptitudes
and Principles I performance scores,
primary sample

N=513

	MAP-RIGHT	MAP-WRONG	MAP-%
EXAM 1	.0909**	-.0565	.0691
EXAM 2	.0366	-.0131	.0196
EXAM 3	.0818*	-.0642	.0705
EXAM 4	.0752*	-.0328	.0466
HW/QUIZ	-.0736*	.0542	-.0725
COMPUTER	-.0548	.0483	-.0635
TOTAL	.0563	-.0274	.0339
	MRT-RIGHT	**MRT-WRONG**	**MRT-%**
EXAM 1	.0276	-.1460***	.1391**
EXAM 2	.0018	-.1135**	.0956**
EXAM 3	.0953**	-.1706****	.1708****
EXAM 4	.0990**	-.1727****	.2023****
HW/QUIZ	-.1328**	.0917**	-.1341**
COMPUTER	-.0581	.0560	-.0512
TOTAL	.0418	-.1387**	.1407**
	SAT-MATH	**SAT-VERBAL**	**GPA**
EXAM 1	.3156****	.1772****	.3757****
EXAM 2	.2056****	.1441**	-.4387****
EXAM 3	.3089****	.1427**	.4289****
EXAM 4	.3641****	.2223****	.4596****
HW/QUIZ	-.1092**	-.1348**	.3616****
COMPUTER	-.0540	-.0964**	.3194****
TOTAL	.2930****	.1544***	.5224****

[**** p = .0001, *** p = .001, ** p < .05, * p < .10]

interpretation that students who are aware of relative aptitude deficiencies may work harder on the two grade components which can be performed outside of class with fewer time constraints than EXAMS.

First tests of an aptitude-performance relationship. To test Hypothesis 2, a two-way MANOVA and a two-way MANCOVA compare performance means among three MAP-RIGHT groups and three MRT-RIGHT groups. MAP-RIGHT groups contain approximately the same number of subjects, as do MRT-RIGHT groups. The categorization scheme is identical to those used for Chi-square analysis of Hypothesis 1.

Another two-way MANOVA and another two-way MANCOVA compare performance means among three levels each of MAP-% and MRT-%. Cut-off scores for each group were chosen from sorted values such that approximately equal numbers of subjects would fall within each of three categories for each variable. MAP-% scores between .923 and .9706 inclusive comprise the middle MAP-% group, and scores above and below these scores comprised the high and low MAP-% categories. For MRT-%, scores between .823 and .925 inclusive comprise the middle group, and scores above and below these scores comprised the high and low MRT-% categories. Subjects were not constrained to be in comparable groups for the two variables.

MANOVA (partial models). Table 15 presents multivariate results of the 2-way MANOVA and univariate p-values for the ANOVA models on VSA levels. MANOVA tests for group differences in the linear combination of dependent variables and so avoids the inflated type 1 error rate which result from separate ANOVAs (Hair, Anderson and Tatham 1987). Multivariate significance for any dependent variable identified below indicates either or both of the following explanations. One, at least one of the dependent variables differs among groups of interest; this difference will be apparent in ANOVA for that variable. Two, the linear combination of dependent variables indicates group differences which are not evident when examining one dependent variable at a time.

In the discussion of results for tests of hypotheses of VSA-performance relationships, variables which are significant in multivariate tests will be considered to indicate a VSA-performance relationship in the course under discussion. Variables which are significant only in univariate tests for a particular performance score will not be considered to indicate a VSA-performance relationship.

In Table 15 Panel A, MAP-RIGHT by MRT-RIGHT analysis shows a VSA-performance relationship for MRT-RIGHT only, indicating that there are performance score differences among the three MRT-RIGHT groups. ANOVA results indicate that MRT-RIGHT is related to EXAM 3. No interaction is found between MAP-RIGHT and MRT-RIGHT.

In Panel B, MAP-% by MRT-% results indicate significance for both MAP-% and MRT-% but not for their interaction. MAP-% is related to EXAMS 1 and 4, while MRT-% is related to all EXAMS and HW/QUIZ. R^2 for all models is low, ranging from less than .01 to slightly more than .06.

Table 15
Results of 2-way MANOVA and ANOVA
of Principles I performance scores
[MAP-RIGHT by MRT-RIGHT and MAP-% by MRT-%]

N = 513 Primary sample

PANEL A: 2-way MANOVA (MAP-RIGHT by MRT-RIGHT) for Prin. I

Wilks' Lambda p-values : MAP-RIGHT .1285; MRT-RIGHT .0048; MAP*MRT-RIGHT .4567

Model	p <	$[R^2]$	Model	p <	$[R^2]$
EXAM 1=		[.0189]	EXAM 2 =		[.0169]
MAP-RIGHT	.3500		MAP-RIGHT	.3325	
MRT-RIGHT	.2896		MRT-RIGHT	.1672	
MAP*MRT-RIGHT	.3593		MAP*MRT-RIGHT	.5977	
EXAM 3=		[.0253]	EXAM 4 =		[.0229]
MAP-RIGHT	.7336		MAP-RIGHT	.4958	
MRT-RIGHT	.0248		MRT-RIGHT	.3722	
MAP*MRT-RIGHT	.3172		MAP*MRT-RIGHT	.1496	
HW/QUIZ =		[.0344]	COMPUTER =		[.0109]
MAP-RIGHT	.0076		MAP-RIGHT	.2409	
MRT-RIGHT	.0730		MRT-RIGHT	.4505	
MAP*MRT-RIGHT	.5043		MAP*MRT-RIGHT	.9922	

PANEL B: 2-way MANOVA (MAP-% by MRT-%) for Principles I

Wilks' Lambda p-values:
MAP-% .0201 MRT-% .0001 MAP*MRT-% .6636

ANOVA Model	p <	$[R^2]$	ANOVA Model	p <	$[R^2]$
EXAM 1=		[.0490]	EXAM 2 =		[.0182]
MAP-%	.0117		MAP-%	.7476	
MRT-%	.0097		MRT-%	.0495	
MAP*MRT-%	.5908		MAP*MRT-%	.8233	
EXAM 3=		[.0479]	EXAM 4 =		[.0602]
MAP-%	.1035		MAP-%	.0510	
MRT-%	.0005		MRT-%	.0001	
MAP*MRT-%	.6657		MAP*MRT-%	.5292	
HW/QUIZ =		[.0266]	COMPUTER =		[.0133]
MAP-%	.7599		MAP-%	.8046	
MRT-%	.0192		MRT-%	.2551	
MAP*MRT-%	.4041		MAP*MRT-%	.5872	

MANCOVA (full models) on MAP and MRT component scores. Table 16 reports multivariate Wilks' Lambda for MANCOVA models and p-values for ANOVA models which use VSA component scores. With the additional variables, model R^2 ranges from .1686 for COMPUTER to .3329 for EXAM 4. MANCOVA significance is found for SAT scores, GPA, and all three demographic variables but not for MAP-RIGHT, MRT-RIGHT, or their interaction.

ANOVA results show that PRIOR and MAJOR are related to all performance scores except HW/QUIZ and COMPUTER. SAT-VERBAL is related to HW/QUIZ and COMPUTER but not to any EXAM. GENDER is significantly related to EXAM 2, HW/QUIZ and COMPUTER, which are the performance scores on which female scores are higher.

In other words, more than 25% of the variance in each EXAM is consistently explained by SAT-MATH, GPA, and PRIOR, with GENDER and MAJOR also contributing to the explained variance. Almost 25% of variance in HW/QUIZ is explained by MAP-RIGHT, SAT-MATH, SAT-VERBAL, GPA, and GENDER. Almost 17% of the variance in COMPUTER grades is explained by SAT-VERBAL, GPA, GENDER and MAJOR.

Table 16
Results of 2-way MANCOVA and ANCOVA
of Principles I performance scores
[MAP-RIGHT by MRT-RIGHT]
(SAT-MATH, SAT-VERBAL, GPA, GENDER, PRIOR, MAJOR)

N = 513 Primary sample

Wilks' Lambda p-values (p < .05 underlined):

MAP-RIGHT	.1564	SAT-MATH	.0001	GENDER	.0001
MRT-RIGHT	.4188	SAT-VERBAL	.0001	PRIOR	.0001
MAP*MRT-RIGHT	.5352	GPA	.0001	MAJOR	.0026

ANCOVA Model	p <	$[R^2]$	ANCOVA Model	p <	$[R^2]$
EXAM 1=		[.2855]	EXAM 2 =		[.2741]
MAP-RIGHT	.6760		MAP-RIGHT	.6586	
MRT-RIGHT	.6222		MRT-RIGHT	.3096	
MAP*MRT-RIGHT	.8186		MAP*MRT-RIGHT	.9287	
SAT-MATH	.0001		SAT-MATH	.0019	
SAT-VERBAL	.1912		SAT-VERBAL	.4156	
GPA	.0001		GPA	.0001	
GENDER	.8840		GENDER	.0127	
PRIOR	.0001		PRIOR	.0008	
MAJOR	.0022		MAJOR	.0001	
EXAM 3=		[.2720]	EXAM 4 =		[.3329]
MAP-RIGHT	.5440		MAP-RIGHT	.7061	
MRT-RIGHT	.1269		MRT-RIGHT	.7336	
MAP*MRT-RIGHT	.4317		MAP*MRT-RIGHT	.5774	
SAT-MATH	.0001		SAT-MATH	.0001	
SAT-VERBAL	.6206		SAT-VERBAL	.1634	
GPA	.0001		GPA	.0001	
GENDER	.0796		GENDER	.0604	
PRIOR	.2072		PRIOR	.0113	
MAJOR	.0008		MAJOR	.0001	
HW/QUIZ =		[.2387]	COMPUTER =		[.1686]
MAP-RIGHT	.0123		MAP-RIGHT	.3660	
MRT-RIGHT	.4485		MRT-RIGHT	.5907	
MAP*MRT-RIGHT	.1547		MAP*MRT-RIGHT	.9664	
SAT-MATH	.0165		SAT-MATH	.2903	
SAT-VERBAL	.0001		SAT-VERBAL	.0008	
GPA	.0001		GPA	.0001	
GENDER	.0075		GENDER	.0046	
PRIOR	.3535		PRIOR	.9440	
MAJOR	.3912		MAJOR	.0546	

MANCOVA (full models) on MAP and MRT percentages. Table 17 reports multivariate Wilks' Lambda p-values for the MANCOVA model and univariate p-values for the ANCOVA models which use VSA percentage scores. MAP-% narrowly misses multivariate significance ($p < .0625$). No interaction is found between MAP-% and MRT-%. SAT scores, GPA and the three demographic variables are all highly significant and reveal the same pattern of univariate relationships described above.

Partial and full regressions--expectations. Also presented are four regression models which include VSA measures as continuous variables. One model includes VSA components only--MAP-RIGHT, MAP-WRONG, MRT-RIGHT, MRT-WRONG. One model includes VSA percentage measures only--MAP-% and MRT-%. The second model includes VSA components and covariates. The fourth includes VSA percentages and covariates. To remove the scale differences of the variables, all aptitude values are standardized (to a distribution with mean = 0 and standard deviation = 1) before use in the models. This facilitates comparison of parameter estimates within each model.

Expectations for the sign of parameter estimates are as follows. EXAMS are expected to be positively related to MAP-RIGHT, MRT-RIGHT, MAP-%, MRT-%, SAT-MATH, SAT-VERBAL and GPA, and negatively related to MAP-WRONG and MRT-WRONG. No theory-based predictions are made regarding relationships to

Table 17
Results of 2-way MANCOVA and ANOVA
of Principles I performance scores
[MAP-% by MRT-%]
(SAT-MATH, SAT-VERBAL, GPA, GENDER, PRIOR, MAJOR)

N = 513 Primary sample

Wilks' Lambda p-values (p < .05 underlined):

MAP-%	.0625	SAT-MATH	.0001	GENDER	.0001
MRT-%	.1509	SAT-VERBAL	.0004	PRIOR	.0001
MAP*MRT-%	.8026	GPA	.0001	MAJOR	.0018

ANOVA Model	p <	$[R^2]$	ANOVA Model	p <	$[R^2]$
EXAM 1=		[.2995]	EXAM 2 =		[.2788]
MAP-%	.0191		MAP-%	.6560	
MRT-%	.2561		MRT-%	.1917	
MAP*MRT-%	.7833		MAP*MRT-%	.6768	
SAT-MATH	.0001		SAT-MATH	.0030	
SAT-VERBAL	.3051		SAT-VERBAL	.3642	
GPA	.0001		GPA	.0001	
GENDER	.4287		GENDER	.0040	
PRIOR	.0001		PRIOR	.0012	
MAJOR	.0017		MAJOR	.0001	
EXAM 3=		[.2733]	EXAM 4 =		[.3407]
MAP-%	.3097		MAP-%	.2100	
MRT-%	.1645		MRT-%	.1149	
MAP*MRT-%	.4960		MAP*MRT-%	.7688	
SAT-MATH	.0001		SAT-MATH	.0001	
SAT-VERBAL	.5354		SAT-VERBAL	.2084	
GPA	.0001		GPA	.0001	
GENDER	.0848		GENDER	.1356	
PRIOR	.3107		PRIOR	.0066	
MAJOR	.0004		MAJOR	.0001	
HW/QUIZ =		[.2380]	COMPUTER =		[.1729]
MAP-%	.7492		MAP-%	.8984	
MRT-%	.0170		MRT-%	.3215	
MAP*MRT-%	.4181		MAP*MRT-%	.5607	
SAT-MATH	.0001		SAT-MATH	.3313	
SAT-VERBAL	.0292		SAT-VERBAL	.0021	
GPA	.0001		GPA	.0001	
GENDER	.0382		GENDER	.0216	
PRIOR	.3704		PRIOR	.9264	
MAJOR	.3843		MAJOR	.0584	

HW/QUIZ and COMPUTER, since these may be tied to motivational variables not included in these models. Correlation analysis, as mentioned above, indicates a negative relationship for HW/QUIZ and COMPUTER with aptitudes except for GPA which is positively related to these two grades.

The Tukey analysis, discussed above, suggests that on EXAMS 1, 2, and 4, the parameter estimate for PRIOR would be positive. Since accounting majors are expected to outperform nonmajors, the relationship of grades to MAJOR is expected to be negative.

The signs of parameter estimates for GENDER are not predicted by theory but are expected to follow the results of gender t-tests, already reported in Table 8. Because females outperform males on EXAM 2, HW/QUIZ and COMPUTER, the GENDER parameter estimates for these performance scores are expected to be negative. Because males outperform females on EXAMS 3 and 4, the GENDER parameter estimates are expected to be positive for these grades. The following discussions generally emphasize only those parameter signs not in keeping with expectations.

Aptitude-performance regression--partial models.[9] Table 18 reports the multivariate and univariate results of regression models in which standardized VSA component scores and then VSA percentage scores, without demographic variables and covariates, are used to predict some of the variance in Principles I grades. R^2 for these partial models for VSA variables only is very low, ranging from about .01 to .04. P-values for all regressions are for tests of the equality of parameter estimates to zero. VSA component scores which show a relationship to performance are MRT-RIGHT and MRT-WRONG. MRT-RIGHT is related to HW/QUIZ only, while MRT-WRONG is explanatory for all exams. For VSA percentage scores, MRT-% is related to the vector of grades and is explanatory for individual EXAMS and HW/QUIZ.

[9]MAP-RIGHT and MRT-RIGHT are normally distributed, but other VSA measures are only marginally so. To approximate normal distributions, VSA variables were transformed (MAP-WRONG and MRT-WRONG with log and MAP-% and MRT-% with arcsin transformations). All regression models reported in this chapter were rerun with substantially similar results. Conclusions concerning hypotheses were identical.

Table 18
Results of regression partial models
of Principles I performance scores
on VSA variables
Primary sample

N = 513

Wilks' Lambda p-values (p < .05 underlined):

MAP-RIGHT	.6974	MRT-RIGHT	.0095
MAP-WRONG	.9302	MRT-WRONG	.0012
MAP-%	.4621	MRT-%	.0001

Univariate parameter estimates for VSA component scores:

EXAM 1 =	80.10****	[.0286]	EXAM 2 =	73.16****	[.0153]
MAP-RIGHT	1.07*		MAP-RIGHT	.74	
MAP-WRONG	.02		MAP-WRONG	.33	
MRT-RIGHT	-.45		MRT-RIGHT	-.63	
MRT-WRONG	-1.74**		MRT-WRONG	-1.82**	
EXAM 3 =	75.67****	[.0350]	EXAM 4 =	94.77****	[.0353]
MAP-RIGHT	.65		MAP-RIGHT	1.38	
MAP-WRONG	-.25		MAP-WRONG	.47	
MRT-RIGHT	.57		MRT-RIGHT	1.08	
MRT-WRONG	-2.17***		MRT-WRONG	-4.00***	
HW/QUIZ =	34.95****	[.0228]	COMPUTER=	54.31****	[.0072]
MAP-RIGHT	-.15		MAP-RIGHT	-.23	
MAP-WRONG	.19		MAP-WRONG	.23	
MRT-RIGHT	-.80**		MRT-RIGHT	-.31	
MRT-WRONG	.43		MRT-WRONG	.35	

Univariate parameter estimates for VSA percentage scores:

EXAM 1 =	80.10****	[.0215]	EXAM 2 =	73.16****	[.0092]
MAP-%	.55		MAP-%	.05	
MRT-%	1.55**		MRT-%	1.42**	
EXAM 3 =	75.67****	[.0309]	EXAM 4 =	94.77****	[.0411]
MAP-%	.60		MAP-%	.31	
MRT-%	2.31***		MRT-%	5.02****	
HW/QUIZ =	34.95****	[.0205]	COMPUTER=	54.31****	[.0057]
MAP-%	-.37		MAP-%	-.48	
MRT-%	-.92**		MRT-%	-.35	

[**** p = .0001, *** p = .001, ** p < .05, * p < .10]

Aptitude-performance regression--full model on MAP and MRT component scores. Table 19 presents the univariate and multivariate results of regression models in which standardized aptitudes, including SAT-MATH, SAT-VERBAL, GPA, and the VSA component scores, are used to explain variance in Principles I scores. Table 20 presents results of a model which replaces VSA component scores with percentage scores. Both models include GENDER, PRIOR, and MAJOR as dichotomous variables. Both tables present multivariate Wilks' Lambda p-values and beta estimates (and significance levels) for univariate analysis.

Table 19, with VSA component models, reports that the only VSA measure related to the grade vector is MRT-WRONG, which has a negative relationship to EXAMS 1, 3, and 4, and a positive relationship to HW/QUIZ. SAT scores, GPA and the three demographic variables are all related to the grade vector. SAT-MATH is significantly related to EXAMS 1, 2, 3

Table 19

Results of regression full models of Principles I performance scores on VSA component scores

N = 513 Primary sample

Wilks' Lambda p-values (p < .05 underlined):

MAP-RIGHT	.9031	MRT-RIGHT	.5029
MAP-WRONG	.7418	MRT-WRONG	.0488
SAT-MATH	.0001	SAT-VERBAL	.0002
GPA	.0001	GENDER	.0001
PRIOR	.0001	MAJOR	.0012

Univariate parameter estimates:

EXAM 1 =	80.46**** [.2865]	EXAM 2 =	77.36**** [.2745]
MAP-RIGHT	.15	MAP-RIGHT	.28
MAP-WRONG	-.18	MAP-WRONG	.34
MRT-RIGHT	-.23	MRT-RIGHT	.45
MRT-WRONG	-.92*	MRT-WRONG	-.93
SAT-MATH	2.87****	SAT-MATH	1.87**
SAT-VERBAL	.63	SAT-VERBAL	.63
GPA	3.46****	GPA	5.69****
GENDER	-.27	GENDER	-3.53**
PRIOR	5.64****	PRIOR	4.09***
MAJOR	-3.47**	MAJOR	-5.42****

EXAM 3 =	77.37**** [.2667]	EXAM 4 =	98.04**** [.3332]
MAP-RIGHT	-.34	MAP-RIGHT	-.62
MAP-WRONG	-.51	MAP-WRONG	.09
MRT-RIGHT	.44	MRT-RIGHT	.81
MRT-WRONG	-.95*	MRT-WRONG	-1.68*
SAT-MATH	2.87****	SAT-MATH	6.06****
SAT-VERBAL	-.35	SAT-VERBAL	1.36
GPA	5.31****	GPA	9.52****
GENDER	2.18*	GENDER	3.30
PRIOR	1.28	PRIOR	5.04**
MAJOR	-4.58***	MAJOR	-9.63****

HW/QUIZ =	36.42**** [.2272]	COMPUTER=	56.73**** [.1698]
MAP-RIGHT	.08	MAP-RIGHT	-.07
MAP-WRONG	.22	MAP-WRONG	.25
MRT-RIGHT	-.12	MRT-RIGHT	.52
MRT-WRONG	.67**	MRT-WRONG	.60
SAT-MATH	-.71**	SAT-MATH	-.38
SAT-VERBAL	-1.38****	SAT-VERBAL	-1.27**
GPA	3.20****	GPA	3.14****
GENDER	-1.49**	GENDER	-2.24**
PRIOR	-.72	PRIOR	-.12
MAJOR	-.50	MAJOR	-1.58*

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

and 4 and HW/QUIZ, whereas SAT-VERBAL is related to HW/QUIZ and COMPUTER, but not to any EXAM. GPA is explanatory for all performance scores. GENDER is explanatory for EXAM 2, HW/QUIZ, and COMPUTER. PRIOR is explanatory for EXAMS 1, 2, and 4, as found previously in Tukey analysis.

In other words, EXAM 1 is negatively related to MRT-WRONG and to MAJOR and positively related to SAT-MATH, GPA, PRIOR. EXAM 2 scores are related to SAT-MATH, GPA, GENDER, PRIOR and MAJOR. EXAM 3 is related to MRT-WRONG and GENDER and significantly to SAT-MATH and GPA and to MAJOR. EXAM 4 is related to MRT-WRONG, SAT-MATH, GPA, PRIOR, and MAJOR. HW/QUIZ is related significantly to MRT-WRONG, SAT-MATH, SAT-VERBAL, GPA, and GENDER. COMPUTER is related to SAT-VERBAL, GPA, GENDER and to MAJOR.

<u>Aptitude-performance regression--full model with MAP and MRT percentages</u>. Table 20, for regression with VSA percentages, reports model R^2 ranging from .1651 (COMPUTER) to .3342 (EXAM 4). Although MAP-% is not related to performance, MRT-% is, consistent with the correlation analysis. MRT-% is positively related to EXAM 4 and negatively related to HW/QUIZ. SAT scores, GPA, GENDER, PRIOR, and MAJOR are all highly significant.

Table 20
Results of regression full models
of Principles I performance scores
on VSA percentage scores

N = 513 Primary sample

Wilks' Lambda p-values (p < .05 underlined):

MAP-%	.7330	GPA	.0001
MRT-%	.0069	GENDER	.0001
SAT-MATH	.0001	PRIOR	.0001
SAT-VERBAL	.0002	MAJOR	.0012

Univariate parameter estimates:

EXAM 1 =	80.61****	[.2844]	EXAM 2 =	77.21****	[.2717]
MAP-%	.33		MAP-%	-.06	
MRT-%	.67		MRT-%	.81	
SAT-MATH	2.90****		SAT-MATH	2.01**	
SAT-VERBAL	.58		SAT-VERBAL	.58	
GPA	3.49****		GPA	5.68****	
GENDER	-.49		GENDER	-3.25**	
PRIOR	5.60****		PRIOR	4.06***	
MAJOR	-3.49**		MAJOR	-5.41****	
EXAM 3 =	77.36****	[.2640]	EXAM 4 =	98.16****	[.3342]
MAP-%	.41		MAP-%	-.22	
MRT-%	.83		MRT-%	2.22**	
SAT-MATH	2.84****		SAT-MATH	5.90****	
SAT-VERBAL	-.33		SAT-VERBAL	1.39	
GPA	5.33****		GPA	9.46****	
GENDER	2.25*		GENDER	3.15	
PRIOR	2.27		PRIOR	4.93***	
MAJOR	-4.56***		MAJOR	-9.63****	
HW/QUIZ =	36.31****	[.2318]	COMPUTER=	56.49****	[.1651]
MAP-%	-.20		MAP-%	-.32	
MRT-%	-.88**		MRT-%	-.27	
SAT-MATH	-.66**		SAT-MATH	-.37	
SAT-VERBAL	-1.38****		SAT-VERBAL	-1.22**	
GPA	3.22****		GPA	3.09****	
GENDER	-1.32**		GENDER	-1.86**	
PRIOR	-.69		PRIOR	-.10	
MAJOR	-.49		MAJOR	-1.55*	

**** p = .0001, *** p = .001, ** p < .05, * p < .10]

Tests of Hypothesis 2--summary. Evidence supports the hypothesized relationship between VSA and performance in Principles I. However, VSA's explanatory power is weak relative to SAT scores and GPA. MANOVA and MANCOVA examine aptitude differences among levels of MAP-RIGHT and MRT-RIGHT and among levels of MAP-% and MRT-%. Partial and full regressions also include MAP-WRONG and MRT-WRONG.

MANOVA (Table 15) results indicate a VSA-performance relationship for MRT-RIGHT (EXAMS), MRT-% (EXAMS and HW/QUIZ), and MAP-% (EXAMS). MANCOVA results (Tables 16 and 17) show a VSA-performance relationship for MAP-% (EXAMS).

Regression without demographic variables or covariates (Table 18) finds a VSA-performance relationship for MRT-RIGHT (HW/QUIZ), MRT-WRONG (EXAMS), and MRT-% (EXAMS and HW/QUIZ). Regression including demographic variables and covariates (Tables 19 and 20) finds multivariate significance for MRT-WRONG (EXAMS and HW/QUIZ) and MRT-% (EXAMS and HW/QUIZ). Neither multivariate nor univariate significance is found for any MAP score.

Results are also found for the non-VSA variables in all models in which they are included. SAT-MATH is related to EXAMS and HW/QUIZ, while SAT-VERBAL is related to HW/QUIZ and COMPUTER. PRIOR and MAJOR are related to EXAMS but not to HW/QUIZ. MAJOR is related to COMPUTER; PRIOR is not. GENDER and GPA are generally related to all dependent variables.

Tests of Hypothesis 3

Hypothesis 3 asserts a gender difference in the aptitude-performance relationship. Reported below are results of *t*-tests for gender comparisons of aptitude variables as well as direct tests of the hypothesis using regression.

Table 21 reports *t*-test comparisons between genders on aptitude variables as well as providing means and standard deviations for each gender. Results show that males score

Table 21
Female and male
aptitude means, standard deviations
and *t*-test comparisons between genders

Variable	Female N=239 Means (s.d.)	Male N=274 Means (s.d.)	*t*-test value
MAP-RIGHT	22.70 (6.36)	25.58 (6.48)	- 5.07****
MAP-WRONG	1.94 (2.75)	1.72 (2.63)	.93
MAP-%	.91 (.13)	.93 (.11)	- 1.65*
MRT-RIGHT	18.93 (6.79)	25.77 (7.84)	-10.58****
MRT-WRONG	4.59 (3.93)	3.48 (3.34)	3.41***
MRT-%	.81 (.14)	.88 (.12)	- 5.97****
SAT-MATH	554.10 (67.46)	587.81 (69.41)	- 5.56****
SAT-VERBAL	473.72 (69.21)	487.85 (67.63)	- 2.33**
GPA	2.53 (.58)	2.48 (.56)	1.05

[**** p = .0001, *** p = .001, ** p < .05, * p < .10]

higher on all aptitude measures except MAP-%, MAP-WRONG, and GPA. Results of nonparametric analysis of means with a Wilcoxon test are comparable to parametric analysis for all variables.

To test Hypothesis 3 on gender differences in the aptitude-performance relationship, general linear models are performed separately for gender sample subsets. These models are identical to those used in analysis of the primary sample for Hypothesis 2 except that GENDER is no longer included as a variable. Standardized aptitudes are also used in these analyses. Tables 22, 23, and 24 present significance levels for multivariate Wilk's Lambda and parameter estimates and associated p-values for univariate analysis.

Aptitude-performance regression by GENDER--partial models. Table 22 reports results of gender models in which VSA scores explain a portion of the variance in performance scores for each gender. This table presents female and male parameter estimates side by side for each grade.

For both genders, MRT-WRONG and MRT-% are related to the grade vector. For females, however, MRT-WRONG is related to HW/QUIZ and to EXAM 2, while for males, MRT-WRONG is related to all EXAMS. For females, MRT-% is related to EXAMS and HW/QUIZ and COMPUTER, while for males MRT-% is only related to EXAMS.

Table 22
Results of female and male regression partial models
of Principles I performance scores
on VSA component and percentage scores
N = 513 (239 females, 274 males)

PANEL A: Wilks' Lambda p-values (p < .05 underlined):

	Female	Male		Female	Male
MAP-RIGHT	.6946	.2841	MRT-RIGHT	.8360	.1688
MAP-WRONG	.6316	.6461	MRT-WRONG	.0110	.0036
MAP-%	.7520	.7548	MRT-%	.0009	.0032

PANEL B: Univariate parameter estimates:

	Female	Male		Female	Male
EXAM 1 =	80.33****	79.95****	EXAM 2=	75.35****	71.09****
MAP-RIGHT	.72	1.52*	MAP-RIGHT	.62	1.36
MAP-WRONG	-.25	.48	MAP-WRONG	-.11	1.04
MRT-RIGHT	.16	-1.00	MRT-RIGHT	.17	.01
MRT-WRONG	-.96	-2.84***	MRT-WRONG	-1.59*	-2.46**
R^2 =	[.0125]	[.0613]	R^2 =	[.0163]	[.0287]
EXAM 3 =	74.44****	76.35****	EXAM 4=	93.57****	96.25****
MAP-RIGHT	.54	.59	MAP-RIGHT	2.44	.30
MAP-WRONG	-.75	.19	MAP-WRONG	1.33	-.22
MRT-RIGHT	-.16	.30	MRT-RIGHT	1.19	.00
MRT-WRONG	-1.39	-3.08***	MRT-WRONG	-2.64	-5.53***
R^2 =	[.0163]	[.0506]	R^2 =	[.0213]	[.0465]
HW/QUIZ =	35.86****	33.72****	COMPUTER=	55.46****	52.64****
MAP-RIGHT	-.51	.45	MAP-RIGHT	-.44	.27
MAP-WRONG	-.12	.67	MAP-WRONG	-.23	.76
MRT-RIGHT	-.60	-.44	MRT-RIGHT	-.56	.69
MRT-WRONG	1.04**	-.51	MRT-WRONG	.47	.01
R^2 =	[.0541]	[.0076]	R^2 =	[.0175]	[.0100]

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

PANEL C: Univariate parameter estimates:

	Female	Male		Female	Male
EXAM 1 =	80.27****	79.66****	EXAM 2 =	75.40****	71.18****
MAP-%	.44	.67	MAP-%	.19	.03
MRT-%	1.10	2.28**	MRT-%	1.86**	2.06**
R^2 =	[.0109]	[.0397]	R^2 =	[.0182]	[.0152]
EXAM 3 =	74.52****	76.22****	EXAM 4 =	93.09****	95.60****
MAP-%	.79	.30	MAP-%	-.12	.67
MRT-%	1.18	3.23***	MRT-%	3.64**	6.03***
R^2 =	[.0115]	[.0484]	R^2 =	[.0212]	[.0527]
HW/QUIZ =	36.04****	33.67****	COMPUTER=	55.68****	52.77****
MAP-%	-.23	-.49	MAP-%	-.13	-.81
MRT-%	-1.26***	.19	MRT-%	-.71*	.93
R^2 =	[.0529]	[.0032]	R^2 =	[.0145]	[.0110]

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

These findings support Hypothesis 3 by indicating a gender difference in the relationships between Principles I performances scores and MRT-WRONG and MRT-%.

Aptitude-performance regression by GENDER--full model on MAP and MRT component scores. Table 23 reports results of gender models in which VSA component scores, demographic variables, and aptitude covariates explain variance in performance scores for each gender. This table presents female and male parameter estimates side by side for each score. R^2 for females ranges from .1439 (COMPUTER) to .3754 (EXAM 4). R^2 for males ranges from .1779 (COMPUTER) to .3079 (EXAM 4).

For females, the grade vector is related to MRT-WRONG as are HW/QUIZ and EXAM 2 and COMPUTER. With the additional variables, males no longer show a relationship between MRT-WRONG and the grade vector. Again, MAP-RIGHT, MAP-WRONG, and MRT-RIGHT are significant for neither gender. Multivariate significance for SAT scores, GPA, and PRIOR is comparable for females and males. Univariate significance, however, does show gender differences for SAT scores. On EXAM 2, for instance, SAT-MATH is significant for males but not for females, while SAT-VERBAL is significant for neither females nor for males. PRIOR is significant on EXAM 4 for females but not for males. MAJOR is significant for females but not for males. The signs of all parameter estimates are as expected.

Table 23

Results of female and male regression full models of Principles I performance scores on VSA component scores

N = 513 (239 females, 274 males)

Wilks' Lambda (p < .05 underlined):

	female	male		female	male
MAP-RIGHT	.6545	.2460	SAT-MATH	.0028	.0001
MAP-WRONG	.4722	.6651	SAT-VERBAL	.0010	.0205
MRT-RIGHT	.8403	.2660	GPA	.0001	.0001
MRT-WRONG	.0052	.1232	PRIOR	.0002	.0009
			MAJOR	.0021	.0935

Univariate parameter estimates [R^2]:

	female	male		female	male
EXAM 1 =	80.29****	80.91****	EXAM 2=	77.28****	73.78****
MAP-RIGHT	.14	.19	MAP-RIGHT	.21	.41
MAP-WRONG	-.33	.00	MAP-WRONG	-.10	.74
MRT-RIGHT	.06	-.50	MRT-RIGHT	.23	.64
MRT-WRONG	-.65	-1.50**	MRT-WRONG	-1.49*	-.59
SAT-MATH	2.82**	2.80****	SAT-MATH	1.36	1.95**
SAT-VERBAL	1.34	-.05	SAT-VERBAL	1.76*	-.13
GPA	3.65****	3.17****	GPA	5.47****	5.96****
PRIOR	5.63***	5.41****	PRIOR	4.99**	3.40**
MAJOR	-2.99*	-4.14**	MAJOR	-5.91**	-4.99**
R^2 =	[.3048]	[.2791]	R^2 =	[.3136]	[.2344]

	female	male		female	male
EXAM 3 =	76.85****	79.84****	EXAM 4=	99.58****	100.49****
MAP-RIGHT	-.10	-.65	MAP-RIGHT	1.53	-2.13
MAP-WRONG	-.91	-.20	MAP-WRONG	1.33	-.76
MRT-RIGHT	-.16	.74	MRT-RIGHT	1.16	.74
MRT-WRONG	-.59	-1.58*	MRT-WRONG	-2.00	-2.14
SAT-MATH	2.82**	2.89***	SAT-MATH	5.10**	6.44****
SAT-VERBAL	-.11	-.70	SAT-VERBAL	3.91**	-.58
GPA	6.07****	4.53****	GPA	8.56****	10.09****
PRIOR	1.04	1.33	PRIOR	8.15**	2.18
MAJOR	-4.07**	-5.14**	MAJOR	-12.90****	-6.71**
R^2 =	[.2939]	[.2344]	R^2 =	[.3754]	[.3079]

	female	male		female	male
HW/QUIZ =	35.86****	35.30****	COMPUTER=	56.26****	54.44****
MAP-RIGHT	-.59	.64	MAP-RIGHT	-.58	.33
MAP-WRONG	-.18	.56	MAP-WRONG	-.32	.63
MRT-RIGHT	-.51	.01	MRT-RIGHT	-.46	1.18*
MRT-WRONG	1.19***	.12	MRT-WRONG	.70*	.80
SAT-MATH	-.58	-.72	SAT-MATH	.07	-.55
SAT-VERBAL	-.98**	-1.85****	SAT-VERB	-1.20**	-1.40**
GPA	2.49****	3.80****	GPA	2.23****	4.00****
PRIOR	-1.30*	-.39	PRIOR	-.92	.61
MAJOR	.30	-1.33	MAJOR	-.94	-2.14
R^2 =	[.2184]	[.2327]	R^2 =	[.1439]	[.1779]

[**** p = .0001, *** p = .001, ** p < .05, * p < .10]

Aptitude-performance regression by GENDER--full model with MAP and MRT percentages. Table 24 reports, for full models with demographic variables, covariates, and VSA percentages, R^2 for females and males ranging from .1421 and .1628 (COMPUTER) to .3713 and .3057 (EXAM 4). MAP-% is significant for neither gender. Results for females show a VSA-performance relationship for MRT-% (HW/QUIZ and COMPUTER). For males, MRT-% is not related to the grade vector at conventional levels. Again, results for SAT scores, GPA, and PRIOR are comparable for females and males, and MAJOR is significant only for females. Gender differences on SAT-VERBAL and PRIOR are comparable to those reported in the model with VSA components and covariates.

Analysis with VSA percentages confirms what is found in the VSA component model, i.e., that MRT-% is no longer significant when SAT-MATH, SAT-VERBAL, GPA, GENDER, PRIOR, and MAJOR are added to the model. The correlation of MRT-% with SAT-MATH, SAT-VERBAL and GPA may well account for this decrease in significance when those aptitudes are added to VSA models. These correlations are greater than MRT-WRONG's correlations with the SAT scores and GPA.

Table 24
Results of female and male regression full models
of Principles I performance scores
on VSA percentage scores
N = 513 (239 females, 274 males)

Wilks' Lambda p-values (p < .05 underlined):

	female	male		female	male
MAP-%	.7570	.9307	GPA	.0001	.0001
MRT-%	.0008	.1013	PRIOR	.0002	.0004
SAT-MATH	.0008	.0001	MAJOR	.0037	.0719
SAT-VERBAL	.0020	.0146			

Univariate parameter estimates [R^2]:

	female	male		female	male
EXAM 1 =	80.28****	80.70****	EXAM 2 =	77.30****	74.12****
MAP-%	.35	.31	MAP-%	.15	-.30
MRT-%	.39	1.19*	MRT-%	1.30	.44
SAT-MATH	2.95***	2.77****	SAT-MATH	1.51	2.16**
SAT-VERBAL	1.24	-.09	SAT-VERBAL	1.63*	-.13
GPA	3.65****	3.25****	GPA	5.45****	5.92****
PRIOR	5.52***	5.55****	PRIOR	4.79**	3.45**
MAJOR	-2.94*	-4.24**	MAJOR	-5.80**	-5.12**
R^2 =	[.3027]	[.2752]	R^2 =	[.3106]	[.2298]
EXAM 3 =	77.01****	79.82****	EXAM 4 =	98.83****	100.02****
MAP-%	.82	-.04	MAP-%	-.26	-.14
MRT-%	.07	1.92**	MRT-%	2.14	2.89**
SAT-MATH	2.96**	2.73***	SAT-MATH	5.44**	5.76****
SAT-VERBAL	-.22	-.65	SAT-VERBAL	3.66**	-.47
GPA	6.08****	4.51****	GPA	8.53****	10.11****
PRIOR	.87	1.26	PRIOR	8.06**	1.97
MAJOR	-4.08**	-5.14**	MAJOR	-12.47****	-6.47**
R^2 =	[.2919]	[.2322]	R^2 =	[.3713]	[.3057]
HW/QUIZ =	36.08****	35.49****	COMPUTER=	56.53****	54.86****
MAP-%	-.15	-.31	MAP-%	-.03	-.68
MRT-%	-1.38****	-.25	MRT-%	-.93**	.42
SAT-MATH	-.68	-.48	SAT-MATH	-.01	-.35
SAT-VERBAL	-.88**	-1.88****	SAT-VERBAL	-1.14**	-1.44**
GPA	2.52****	3.77****	GPA	2.24****	3.72****
PRIOR	-1.23*	-.30	PRIOR	-.92	.50
MAJOR	.13	-1.44	MAJOR	-1.09	-2.25
R^2 =	[.2172]	[.2292]	R^2 =	[.1421]	[.1628]

**** p = .0001, *** p = .001, ** p < .05, * p < .10

Tests of Hypothesis 3--summary. Evidence indicates a gender difference in the relationship between Principles I scores and VSA, although this difference appears to be generally limited to the aspect of VSA measured by the MRT. MRT-WRONG and MRT-% tend to be related to HW/QUIZ and COMPUTER performance for females and EXAM performance for males. These findings support the hypothesized gender difference in the relationship between VSA and performance in Principles I.

Tests of Hypothesis 4

Hypothesis 4 predicts a VSA difference between those students who continue to Principles II (continuers) and those students who do not continue to Principles II (noncontinuers) in the semester following Principles I. The sample used for this analysis consists of the 470 subjects who passed Principles I and were eligible to continue to Principles II. The tests of this hypothesis include multiple discriminant analysis, *t*-tests, and Chi-square.

Table 25 reports aptitude means for each group and the results of *t*-tests between group means. *T*-tests on group differences indicate that continuers have higher MRT-%, SAT-MATH and GPA than do noncontinuers, as well as having scored higher on the total points in Principles I. The p-value for a VSA-performance relationship for MRT-WRONG is .0512, which

Table 25
Comparison of aptitudes and Principles I total across continuance status

	N=340	N=130	
	Means (s.d.)		t-test
Variable	Continuers	Noncontinuers	value
MAP-RIGHT	24.43 (6.79)	23.89 (6.19)	.79
MAP-WRONG	1.81 (2.73)	1.82 (2.78)	- .04
MAP-%	92% (12%)	93% (11%)	- .11
MRT-RIGHT	22.92 (8.22)	21.85 (7.99)	1.28
MRT-WRONG	3.75 (3.57)	4.47 (3.58)	-1.96*
MRT-%	85% (13%)	83% (13%)	2.02**
SAT-MATH	581.05 (71.07)	557.23 (66.47)	3.31***
SAT-VERBAL	484.23 (65.82)	479.00 (77.64)	.68
GPA	2.61 (.55)	2.35 (.58)	4.56****
Principles I total grade	438.54 (46.93)	386.98 (47.57)	10.63****

**** p = .0001 *** p < .001 ** p < .05 * p < .10

narrowly misses conventional significance of .05. The finding for MRT-%, however, clearly supports the hypothesized relationship between VSA and Principles I scores.

Discriminant analysis. T-tests examine one variable at a time for differences between groups. Multiple discriminant analysis uses a linear combination of the independent variables to examine group differences and is an appropriate technique to use in testing hypotheses involving a categorical dependent variable and several metric independent variables (Hair, Anderson and Tatham 1987). Table 26 reports the results of multiple discriminant analysis of aptitude variables to predict subjects' identities as either continuers or noncontinuers. The analysis sample contains 60% of the consenting records, while the test (holdout) sample contains the remaining records. The proportional chance criterion, against which hit ratios are measured, is .5918. Hit ratios must exceed this number for the analysis to be helpful.

Ten models are examined, five with VSA component scores and five with VSA percentage scores. For each set, models are as follows: (1) VSA scores, (2) VSA scores and SAT-MATH, (3) VSA scores and SAT-VERBAL, (4) VSA scores and Principles I TOTAL1, and (5) VSA, SAT scores, and TOTAL1. The hit ratio for all ten models exceed the proportional chance rate. The VSA-only models, however, assign all individuals to the group of continuers and thereby achieve a maximum chance rate. The discriminant models are not valid discriminators between the groups of interest.

Chi-square. Table 26 Panel B reports results of Chi-square tests for independence of continuance rates and VSA level. Cut-off points for VSA groups were recalculated for the subjects eligible to continue to Principles II but were substantially the same as used for other categorical analyses in this study. Because Chi-square is appropriate for frequencies rather than percentages, only the VSA components are tested. Panel B reports that the MAP-WRONG and MRT-WRONG measures are not independent of VSA levels, at $p < .05$. Subjects with high MAP-WRONG or MRT-WRONG continue to Principles II at a lower rate than expected, given the

Table 26, Panel A
Results of discriminant analyses of standardized aptitudes to differentiate continuers and noncontinuers
Proportional Chance = .5918 [10]
Classification Matrices for Holdout Sample (N = 195)

Variables in Models	Identity	Classification -Cont-	-Non-	Hit Ratio[11]
Model A: VSA component scores	Cont.	142	0	
	Non.	53	0	.7282
Model B: VSA comp. scores and SAT-MATH	Cont.	141	1	
	Non.	52	1	.7282
Model C: VSA comp. scores and SAT-VERBAL	Cont.	142	0	
	Non.	53	0	.7282
Model D: VSA comp. scores and Prin. I. TOTAL	Cont.	126	16	
	Non.	27	26	.7795
Model E: VSA comp. scores SAT-MATH, SAT-VERBAL Prin. I. TOTAL	Cont.	127	15	
	Non.	27	26	.7846
Model F: VSA percentage variables	Cont.	142	0	
	Non.	53	0	.7282
Model G: VSA % variables and SAT-MATH	Cont.	140	2	
	Non.	53	0	.7179
Model H: VSA % variables and SAT-VERBAL	Cont.	142	0	
	Non.	53	0	.7282
Model I: VSA % variables and Prin. I. TOTAL	Cont.	128	14	
	Non.	27	26	.7897
Model J: MAP-%, MRT-%,SAT-MATH, SAT-VERBAL, TOTAL	Cont.	126	16	
	Non.	26	27	.7846

[10] The proportional chance ratio (C^{Pro}) is calculated as the sum of the square of the proportion of completers $(.7143)^2$ plus the square of the proportion of droppers $(.2857)^2 = .5918$.

[11] Hit ratio is calculated as the number correctly classified divided by the number classified. This is compared to the proportional chance ratio.

Table 26, Panel B
Results of Chi-square tests of independence
between VSA components levels and continuance status

Variable	Continuers actual	Continuers expected	Noncontinuers actual	Noncontinuers expected	X^2
MAP-WRONG					
low	100	107.96	49	41.038	
medium	111	96.34	22	36.63	
high	131	137.67	59	52.331	11.371**
MRT-RIGHT					
low	117	115.21	42	43.792	
medium	107	117.38	55	44.619	
high	118	109.41	33	41.589	5.883*
MRT-WRONG					
low	157	147.09	46	55.911	
medium	115	113.76	42	43.242	
high	70	81.153	42	30.847	8.039**

**** p = .0001, *** p < .001, ** p < .05, * p < .10

proportion of continuers to noncontinuers in the sample. This finding is as expected. However, subjects with low MAP-WRONG also continued less than expected while subjects with low MRT-WRONG continued more than expected. In fact, low and high MAP-WRONG scorers continue less than expected, while medium MAP-WRONG scorers continued at higher rates than expected. The explanation of this anomaly is not apparent. The p-value associated with MRT-RIGHT is .053 and narrowly misses conventional significance.

Tests of Hypothesis 4--summary. Two of three tests indicated support for the hypothesized relationship between VSA and continuance to Principles II. Also, tests of VSA component scores indicate that continuing is not independent of MAP-WRONG and MRT-WRONG. In tests of VSA percentage scores, continuers have higher MRT-% scores than noncontinuers, which indicates that continuers and noncontinuers differ in their VSA accuracy.

Tests of Hypothesis 5

Hypothesis 5 asserts an aptitude-treatment interaction, or difference in the aptitude-performance relationship in different courses. To test this hypothesis, general linear models are used in two approaches. First, models are reported in which the independent variables are used to explain variance in Principles II grades for the group of subjects who continued to Principles II (N = 342). The parameter estimates resulting from the models are presented side by side with recalculated (with N = 342) estimates for Principles I. It should be noted that some grades for Principles I and II are on a slightly different scale. EXAM 2 is 100 for Principles I and 150 for Principles 2; HW/QUIZ is 40 and 60; and COMPUTER is worth 50 and 40. This should be kept in mind when comparing regression results for Principles I and II.

Second, for tertiary samples of subjects continuing to either Cost or Intermediate, for instance, parameter estimates are compared for models on the exams in each course. Analysis excludes the HW/QUIZ and COMPUTER measures from Principles I and II because data for Cost and Intermediate are only available for the first two exams. So, for subjects continuing to Cost, parameters for three models (Principles I, Principles II, Cost) are compared. Three comparable models are reported for those students who continued to Intermediate. For all models, standardized aptitudes are used.

For each tertiary sample, the analysis in this second approach compares the relationship between aptitude and performance in the courses in which they enrolled. The sample sizes for these analyses are greatly reduced from that for Principles I.

Descriptive statistics--secondary and tertiary samples. Before presenting tests of Hypothesis 5, descriptive statistics are presented for aptitudes for the secondary and one of the tertiary samples in Table 27. Results for the secondary sample are similar to those results for the primary sample. The tertiary sample described in Tables 27 and 28 is the sample of those original consenting completers of Principles II who continued to both Intermediate and Cost (N = 71). This tertiary sample is analyzed in a separate subsection, after analysis of subjects who continued to

Table 27
Aptitude descriptive statistics
for secondary and tertiary samples

PANEL A--secondary sample, N = 342

variable	mean	standard deviation	minimum	maximum
MAP-RIGHT	24.44	6.80	0.00	40.00
MAP-WRONG	1.82	2.74	0.00	23.00
MAP-%	92%	12%	0%	100%
MRT-RIGHT	22.89	8.24	3.00	40.00
MRT-WRONG	3.78	3.57	0.00	26.00
MRT-%	85%	13%	35%	100%
SAT-MATH	580.79	71.13	390.00	800.00
SAT-VERBAL	483.88	65.62	310.00	680.00
GPA	2.61	.55	1.30	4.00

PANEL B--tertiary sample, N = 71

variable	mean	standard deviation	minimum	maximum
MAP-RIGHT	25.32	6.93	1.00	40.00
MAP-WRONG	1.76	3.29	0.00	23.00
MAP-%	93%	14%	5%	100%
MRT-RIGHT	22.17	8.17	3.00	40.00
MRT-WRONG	3.44	3.91	0.00	26.00
MRT-%	86%	14%	23%	100%
SAT-MATH	597.04	70.88	430.00	750.00
SAT-VERBAL	493.38	68.58	350.00	650.00
GPA	2.86	.55	1.57	4.00

Intermediate (N = 94) and analysis of subjects who continued to Cost (N = 75). Results of all three tertiary samples are substantially similar.

Table 28 presents, for secondary and tertiary samples, _t_-tests of gender differences in aptitudes. Panel A reports female and male aptitude means, standard deviations, and t-test values for gender differences for the secondary sample. Panel B reports the same items for the tertiary sample. Panel C provides a comparison of these results with the results for the primary sample.

All other tables denote the presence of significant parameters in models by gender. Next to each parameter listed for the two-gender sample, symbols (f, m, F, M) signify that the parameter for that variable is also significant at $p < .10$ or $p < .05$ for females or males. Absence of such symbols in these tables indicates that the parameter for that variable is significant in neither gender model. The symbols are provided to suggest whether females, males, or both are driving the results of the two-gender models.

Table 28
Comparison of *t*-tests between genders across samples
and
aptitude means, standard deviations and *t*-test results across genders for secondary and tertiary samples

PANEL A: secondary sample details (150 females, 190 males)

Variable	Means (s.d.) Female		Male		*t*-test value
MAP-RIGHT	22.93	(6.34)	25.63	(6.71)	-3.69***
MAP-WRONG	1.87	(2.57)	1.78	(2.86)	.32
MAP-%	92%	(12%)	93%	(12%)	- .88
MRT-RIGHT	18.85	(6.88)	26.08	(7.83)	-8.90****
MRT-WRONG	4.51	(4.11)	3.20	(2.98)	3.29**
MRT-%	81%	(15%)	89%	(11%)	-5.59****
SAT-MATH	564.80	(69.64)	593.42	(69.92)	-3.75***
SAT-VERBAL	479.60	(70.43)	487.26	(61.54)	-1.07
GPA	2.65	(.56)	2.58	(.54)	1.13

PANEL B: tertiary sample details (37 females, 34 males)

Variable	Means (s.d.) Female		Male		*t*-test value
MAP-RIGHT	23.70	(6.75)	27.09	(6.78)	-2.11**
MAP-WRONG	1.68	(3.42)	1.85	(3.19)	- .23
MAP-%	92%	(17%)	93%	(11%)	- .23
MRT-RIGHT	20.35	(6.69)	24.15	(9.21)	-2.00*
MRT-WRONG	3.97	(4.97)	2.85	(2.16)	1.24
MRT-%	85%	(15%)	87%	(14%)	- .73
SAT-MATH	593.24	(70.00)	601.18	(72.65)	- .47
SAT-VERBAL	498.38	(68.74)	487.94	(69.00)	.64
GPA	2.95	(.60)	2.78	(.48)	1.25

[**** p = .0001, *** p = .001, ** p < .05, * p < .10]

PANEL C: Comparison of *t*-test results across samples

Variable	Sample	Males Higher	Females Higher	No Difference
MAP-RIGHT	primary	x		
	secondary	x		
	tertiary	x		
MAP-WRONG	primary			x
	secondary			x
	tertiary			x
MAP-%	primary	x		
	secondary			x
	tertiary			x
MRT-RIGHT	primary	x		
	secondary	x		
	tertiary			x
MRT-WRONG	primary		x	
	secondary		x	
	tertiary			x
MRT-%	primary	x		
	secondary	x		
	tertiary			x
SAT-MATH	primary	x		
	secondary	x		
	tertiary			x
SAT-VERBAL	primary	x		
	secondary			x
	tertiary			x
GPA	primary			x
	secondary			x
	tertiary			x

First approach: regression comparison across two courses--Principles I and II. Tables 29, 30, and 31 report regression estimates for Principles II next to new estimates for Principles I, both based on the secondary sample (N = 342). This permits isolating parameter differences attributable to course differences, since the sample is identical for both sets of estimates presented in Tables 29 through 31. Bracketed comments about separate regressions by gender follow immediately after discussion of each variable in the two-gender sample.

Partial models. Table 29 reports regression results on partial models, for VSA measures only, on EXAMS, HW/QUIZ and COMPUTER.[12] For the secondary sample, Principles II scores but not Principles I scores are related to MAP-RIGHT [EXAMS] for the sample as a whole [and for males]. MRT-RIGHT is related to Principles I for the sample as a whole [and for females] and for Principles II for the sample as a whole [but for neither gender when females and males are analyzed separately]. Univariate relationships for the sample as a whole for MRT-RIGHT are as follows: MRT-RIGHT and Principles I HW/QUIZ, MRT-RIGHT and Principles II vector of scores.

[12]The Principles I parameter estimates in Table 29 (secondary sample, N = 342) can be compared to the estimates in Table 18 (primary sample, N = 513). With the smaller sample size, MRT-W is no longer significant.

Neither MAP-WRONG nor MRT-WRONG is significant for either course [or for either gender in separate analysis].

Table 29 indicates that MRT-% is related to performance scores for both Principles I and II in the two-gender sample [but is significant for females only in Principles I and for males only in Principles II]. Associated univariate relationships for the sample as a whole are MRT-% with Principles I EXAMS and HW/QUIZ [and for females] and MRT-% with Principles II EXAMS [and for both genders].

Table 29
Results of regression partial models
of Principles I and II performance scores
on VSA variables, secondary sample[13]

N = 337

Wilks' Lambda p-values (p < .05 underlined):

	Prin I	Prin II		Prin I	Prin II
MAP-RIGHT	.3531	.0173 M	MAP-%	.7049	.7072
MAP-WRONG	.9994	.2046	MRT-%	.0057 F	.0048 M
MRT-RIGHT	.0167 F	.0476			
MRT-WRONG	.1653	.1084			

Univariate parameter estimates [R^2]:

	Prin I	Prin II		Prin I	Prin II
EXAM 1=	83.64****FM	74.69****FM	EX 2=	78.22****FM	109.18****FM
MAP-RT	.89 M	-.72 M		.05	.34
MAP-WR	-.14	-.80 M		-.20	-1.50
MRT-RT	-.21 F	-.40		-.59	-.40
MRT-WR	-1.01* M	-1.62** F		-1.45** FM	-1.61
R^2 =	[.0209]	[.0220]	R^2 =	[.0153]	[.0137]
EXAM 3=	80.75****FM	87.84****FM	EX4=	103.60****FM	101.96****FM
MAP-RT	.75	-.08		1.44	2.71** F
MAP-WR	-.02	-.45		.23	1.10
MRT-RT	.08	-.38		.51	1.63
MRT-WR	-1.51** F	-1.38** FM		-2.63** F	-2.67**
R^2 =	[.0302]	[.0287]	R^2 =	[.0254]	[.0490]
HW/QU=	36.50****FM	44.73****FM	COMP=	56.09****FM	36.42****FM
MAP-RT	.31	-.24		-.30	-.53
MAP-WR	-.02	-.13		-.07	-.34
MRT-RT	-.96** F	-.78 F		-.07 M	.10
MRT-WR	.14	.00		.18	-.21
R^2 =	[.0320]	[.0112]	R^2 =	[.0042]	[.0087]

**** p = .0001, *** p = .001, ** p < .05, * p < .10

Table continued next page.

[13]Females: F = sig. (p < .05) f = sig. (p < .10)
Males: M = sig. (p < .05) m = sig. (p < .10)

Table 29, continued

EXAM 1=	83.64****FM		74.69****FM		EX 2=	78.22****FM		109.18****FM	
MAP-%	.73		.41			.23		1.80	
MRT-%	.88*		1.29*	F		.90	F	1.41	
R^2 =	[.0163]		[.0134]		R^2 =	[.0068]		[.0141]	
EXAM 3=	80.75****FM		87.84****FM		EX4=	103.60****FM		101.96****FM	
MAP-%	.58		.36			.65		.56	
MRT-%	1.49**	F	1.16**	FM		3.48**	F	4.15***	FM
R^2 =	[.0270]		[.0224]		R^2 =	[.0315]		[.0431]	
HW/QU=	36.50****FM		44.73****FM		COMP=	56.09****FM		36.42****FM	
MAP-%	.08		-.17			-.21		-.08	
MRT-%	-.66**	F	-.25			-.27		.30	M
R^2 =	[.0147]		[.0016]		R^2 =	[.0040]		[.0029]	

**** p = .0001, *** p = .001, ** p < .05, * p < .10

<u>Full models--MAP and MRT components</u>. Table 30 reports, for the same secondary sample[14], results of regressions for Principles I and II grades which include SAT-MATH, SAT-VERBAL, GPA, GENDER, PRIOR and MAJOR. In full models for the secondary sample, no VSA measure is related to performance in either Principles I or Principles II for the two-gender sample. Principles I scores are related to SAT-MATH [EXAMS], SAT-VERBAL [HW/QUIZ and COMPUTER], and GPA [all measures], while Principles II scores are related to SAT-MATH [EXAMS and HW/QUIZ] and GPA [all measures]. Principles I but not

[14]New Principles I parameter estimates in Table 30 (for the secondary sample, N = 342) can be compared to the original estimates in Table 19 (primary sample, N = 513). With the smaller sample size, MRT-W is no longer related to Principles I performance.

Principles II scores are related to GENDER [EXAM 2, HW/QUIZ and COMPUTER], PRIOR [EXAMS], and MAJOR [EXAMS]. Of course, PRIOR is a measure taken at the start of Principles I. PRIOR's significance in Principles I but not Principles II is reasonable. By the start of Principles II, all subjects have had prior exposure to accounting by virtue of having just been in Principles I.

In summary, for the two-gender sample, course differences are apparent, but they do not involve VSA.

In examining regressions by gender, however, female Principles I but not Principles II scores are related to MRT-RIGHT [HW/QUIZ]. Although GPA is significant for both genders in both courses, SAT-MATH is significant in both courses only for males, while SAT-VERBAL is significant in both courses only for females. So female Principles I scores are explained by MRT-RIGHT, SAT-VERBAL, GPA, and both demographic variables, while male Principles I scores are explained by SAT-MATH, SAT-VERBAL, GPA and both demographic variables. For females, Principles II scores are related to SAT-VERBAL and GPA, while for males these scores are related to SAT-MATH and GPA. In summary, a course difference in the VSA-performance relationship is observable for females but not for males, while a gender difference in other aptitudes (mathematical and verbal) is apparent in both courses.

Table 30
Results of regression full models
of Principles I and II performance scores
on VSA component scores, secondary sample[15]

N = 337

PANEL A: Wilks' Lambda p-values (p <.05 underlined)

	Prin I	Prin II		Prin I	Prin II
MAP-RIGHT	.7865	.1333	SAT-VERB	.0069 FM	.0749 F
MAP-WRONG	.9977	.2888	GPA	.0001 FM	.0001 FM
MRT-RIGHT	.1197 F	.1459	GENDER	.0001	.3829
MRT-WRONG	.8052	.6020	PRIOR	.0014 FM	.8337
SAT-MATH	.0001 M	.0001 M	MAJOR	.0320 FM	.2447

PANEL B: Univariate parameter estimates:

	Prin I	Prin II		Prin I	Prin II
EXAM 1=	84.48****FM	75.36****FM	EX 2=	81.79****FM	111.57****FM
MAP-RT	.16	-1.40* M		-.34	-1.26 M
MAP-WR	-.30	-.86 M		-.17	-1.71 m
MRT-RT	.26	-.79*		.39	-.24
MRT-WR	-.26	-.27		-.66	.51
SAT-M	2.29****FM	2.44*** fM		1.93** FM	3.49** M
SAT-V	.40	.05 Fm		-.48	.06 F
GPA	2.64****FM	4.85****FM		4.13****FM	9.16****FM
GENDER	-1.88*	2.06**		-3.89**	3.91**
PRIOR	3.93****FM	.50		2.65** F	1.35
MAJOR	-2.12* M	-2.62* m		-3.36** M	-6.14** FM
R^2 =	[.2257]	[.2543]	R^2 =	[.2422]	[.2735]
EX 3=	82.39****FM	88.57****FM	EX4=	106.83****FM	102.10****FM
MAP-RT	-.07	-.43		-.31	.92 F
MAP-WR	-.28	-.33		-.18	.91
MRT-RT	-.13	-.25		.25	1.19 m
MRT-WR	-.61	-.67		-.44	-.12
SAT-M	2.26*** FM	.54		5.60****FM	4.67**** M
SAT-V	-.46	.48* f		.12	1.78* F
GPA	3.32****FM	3.50****FM		7.44****FM	8.37****FM
GENDER	1.75	.49		1.95	3.47
PRIOR	.10	-.12		1.53 F	2.01
MAJOR	-3.66** FM	-1.35		-6.83** F	-4.11* F
R^2 =	[.2230]	[.2259]	R^2 =	[.2812]	[.3290]

**** p = .0001, *** p = .001, ** p < .05, * p < .10

Panel B continued next page.

[15]Females: F = sig. (p < .05) f = sig. (p < .10)
Males: M = sig. (p < .05) m = sig. (p < .10)

PANEL B, continued:

Univariate parameter estimates:

	Prin I	Prin II		Prin I	Prin II
HW/QU=	37.91****FM	45.14****FM	COMP=	57.87****FM	36.94****FM
MAP-RT	.31	-.08		-.32	-.50
MAP-WR	.09	.10		-.02	-.26
MRT-RT	-.36 F	-.17 f		.72** M	.51 M
MRT-WR	.31	.22		.20	-.08
SAT-M	-.31	-1.22** m		-.22	-.36
SAT-V	-1.09*** FM	-1.08** F		-1.08** FM	-.50 m
GPA	1.86****FM	2.61****FM		1.37****FM	1.20**** M
GENDER	-1.48**	-.88		-2.28***	-1.00
PRIOR	-.43	-.15		.69	.46
MAJOR	-.60	.21		-1.17* M	-.21
R^2 =	[.1900]	[.1270]	R^2 =	[.1422]	[.0731]

**** p = .0001, *** p = .001, ** p < .05, * p < .10

Full model--MAP and MRT percentages. Table 31 reports, for these same subjects, the results of models identical to those just discussed except that VSA percentages are used instead of VSA component scores. As is the case for VSA components, no VSA percentage is related to either course for the two-gender sample. The relationships of the non-VSA variables are as reported in VSA component models.

In examining separate regressions by gender, however, female Principles I but not Principles II scores are related to MRT-%. Again, a course difference in the VSA-performance relationship is identified, but only for females.

The VSA relationships already reported in Table 29 (partial models) are generally not observed when the full

Table 31
Results of regression full models of Principles I and II performance scores on VSA percentages, secondary sample[16]
N = 337

PANEL A: Wilks' Lambda p-value (p <.05 underlined)

	Prin I	Prin II		Prin I	Prin II
MAP-%	.8293	.7857	GPA	.0001 FM	.0001 FM
MRT-%	.4452 F	.3636	GENDER	.0001	.1862
SAT-MATH	.0001 M	.0001 M	PRIOR	.0014 FM	.7442
SAT-VERBAL	.0054 FM	.1049 F	MAJOR	.0309 FM	.2462

PANEL B: Univariate parameter estimates:

	Prin I	Prin II		Prin I	Prin II
EX 1=	84.38****FM	76.00****FM	EX 2=	81.75****FM	112.08****FM
MAP-%	.52	.13		.04	1.24
MRT-%	.25	-.27		.49	-.90
SAT-M	2.34****FM	2.05** fm		1.91** FM	3.13** M
SAT-V	.40	.18 f		-.45	.24 F
GPA	2.63****FM	4.96****FM		4.14****FM	9.23****FM
GENDER	-1.69*	1.20		-3.74**	1.96
PRIOR	3.94****FM	.16		2.54** F	1.01
MAJOR	-2.13* M	-2.61*		-3.34** M	-6.10** Fm
R^2 =	[.2258]	[.2397]	R^2=	[.2390]	[.2723]
EX 3=	82.53****FM	88.85****FM	EX4=	107.00****FM	101.53****FM
MAP-%	.40	.12		.18	-.10
MRT-%	.36	.46		.97	.94
SAT-M	2.25****FM	.42		5.47****FM	4.94**** M
SAT-V	-.46	.50 F		.14	1.69 F
GPA	3.34****FM	3.53****FM		7.33****FM	8.18****FM
GENDER	1.60	.12		1.69	4.30**
PRIOR	.03	-.24		1.54 F	2.31
MAJOR	-3.70** FM	-1.37		-6.87** F	-4.14* F
R^2 =	[.2220]	[.2204]	R^2 =	[.2825]	[.3254]

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

Panel B continued, next page.

[16]Females: F = sig. (p < .05) f = sig. (p < .10)
Males: M = sig. (p < .05) m = sig. (p < .10)

Panel B, continued

HW/QU=37.89****FM	45.21****FM	COMP=57.62****FM	36.89****FM
MAP-%	-.01	-.27	-.25	-.13
MRT-%	-.58* F	-.15	.05	.38
SAT-M	-.23	-1.25** m	-.25	-.47**
SAT-V	-1.12*** FM	-1.08** F	-1.01** FM	-.42 m
GPA	1.92****FM	2.61****FM	1.33****FM	1.15*** M
GENDER	-1.47**	-1.03	-1.88**	-.92
PRIOR	-.41	-.14	.66	.39
MAJOR	-.59	.22	-1.10 M	-.17
R^2 =	[.1910]	[.1268]	R^2 = [.1324]	[.0670]

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

models are run, as reported in Table 30 and 31. This is presumably due to multicollinearity among VSA and non-VSA variables. Although correlations between VSA measures and other aptitude variables are generally weak, MRT measures are correlated with SAT-MATH above .20. GPA may well reduce the unique explanatory power of VSA measures; it is reasonable to suppose that GPA has captured some of the effect of VSA on performance in coursework in general. Given these assumptions, regressions with VSA alone (partial models) might be expected to show more significance for VSA measures than would full models, and they do.

<u>Second approach: regression comparisons across three courses</u>. Results of the second approach to testing Hypothesis 5 are reported in Tables 32 through 37: Tables 32 through 34 for Principles I, II and Intermediate models for the tertiary sample of subjects taking Intermediate, and Tables 35 through 37 for the tertiary sample of subjects taking Cost. Table

formats present information for three courses side-by-side. Only EXAMS are analyzed because Intermediate and Cost performance scores do not include either HW/QUIZ or COMPUTER.

Intermediate--partial models. Table 32 reports results of partial models which use VSA measures alone to explain variance in the exams of Principles I and II and Intermediate. R^2 is low for both exams for all courses, consistent with other regressions already reported which include only VSA scores as independent measures. For this reduced sample (N = 94), no VSA measure was significantly related to any EXAM in Principles I, II, or Intermediate.

Table 32
Results of regression partial models
of Principles I, II and Intermediate performance scores
on VSA component or percentage scores, tertiary sample[17]
N = 94 (53 females, 41 males)

PANEL A (VSA components):

Wilks' Lambda p-values:

	Prin I	Prin II	Inter
MAP-RIGHT	.7035	.9313	.1458
MAP-WRONG	.9338	.6971	.8010
MRT-RIGHT	.2373	.4576	.8999
MRT-WRONG	.3214	.7401	.7105

Univariate parameter estimates:

	Prin I	Prin II	Inter
Exam 1 =	88.10****F M	80.02****F M	77.17****F M
MAP-RIGHT	.30	.94 F M	2.95* F
MAP-WRONG	-.19	.44	.61
MRT-RIGHT	.77 F	-.47	-.48
MRT-WRONG	.59	-.83	-.70
R^2 =	[.0150]	[.0087]	[.0492]
Exam 2 =	84.56****F M	120.68****F M	67.00****F M
MAP-RIGHT	-.11	.99	1.31 F
MAP-RIGHT	.64	-.59	-.22
MRT-RIGHT	.49	-2.06	-.75
MRT-WRONG	.00	-1.18	-1.27 M
R^2 =	[.0075]	[.0150]	[.0171]
Exam 3 =	84.51****F M	91.44****F M	
MAP-RIGHT	1.32	.78	
MAP-WRONG	-.23	.87	
MRT-RIGHT	-1.17	-.14	
MRT-WRONG	-1.46	-.75	
R^2 =	[.0430]	[.0203]	
Exam 4 =	114.97****F M	111.04****F M	
MAP-RIGHT	-.31	1.80	
MAP-RIGHT	.13	1.74	
MRT-RIGHT	2.20	1.79	
MRT-WRONG	.55	-2.81	
R^2 =	[.0138]	[.0557]	

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

[17]Females: F = sig. (p < .05) f = sig. (p < .10)
Males: M = sig. (p < .05) m = sig. (p < .10)

PANEL B (VSA percentages):

Wilks' Lambda p-values:

	Prin I	Prin II	Inter
MAP-%	.4367	.6380	.5151
MRT-%	.4987	.0942	.2332

Univariate parameter estimates:

	Prin I	Prin II	Inter
Exam 1 =	88.10****F M	80.02****F M	77.17****F M
MAP-%	.42	.25	1.37
MRT-%	.10	.88	1.10
R^2 =	[.0030]	[.0082]	[.0264]
Exam 2 =	84.56****F M	120.68****F M	67.00****F M
MAP-%	-.48	1.89	.76
MRT-%	.48	.19	2.39* M
R^2 =	[.0043]	[.0113]	[.0377]
Exam 3 =	84.51****F M	91.44****F M	
MAP-%	1.11	-.26	
MRT-%	1.07 F	1.03	
R^2 =	[.0322]	[.0226]	
Exam 4 =	114.97****F M	111.04****F M	
MAP-%	-.32	-.28	
MRT-%	2.64	4.59** F	
R^2 =	[.0229]	[.0644]	

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

Intermediate--full model with MAP and MRT components. Table 33 reports Wilks' Lambda p-values, parameter estimates and associated p-values for full models for EXAMS for all three courses. For this reduced sample, no VSA scores are related to Principles I or II scores. For Intermediate, MAP-RIGHT narrowly misses significance at conventional levels (p < .0646). Neither SAT-MATH nor SAT-VERBAL is related to any measure in any course. GPA is related to EXAM performance in all courses. PRIOR is significant for

Table 33
Results of regression full models of Principles I, II and Intermediate performance scores on VSA component scores, tertiary sample[18]

N = 94 (53 females, 41 males)

Wilks' Lambda p-values (p < .05 underlined):

	Prin I	Prin II	Inter
MAP-RIGHT	.7309	.9118	.0646
MAP-WRONG	.7180	.6129	.4636
MRT-RIGHT	.2159	.3324	.3722
MRT-WRONG	.4110	.8792	.9611
SAT-MATH	.2085	.5489	.9176
SAT-VERBAL	.3513	.8342	.2426
GPA	.0001 F M	.0001 F M	.0001 F M
GENDER	.4890	.2656	.2292
PRIOR	.0445 F	.5746	.6217
MAJOR	.3112 F	.6460	.0637

Univariate parameter estimates:

	Prin I	Prin II	Inter
Exam 1 =	85.62****F M	78.70****F M	74.69****F M
MAP-RIGHT	.64	.21 f M	3.63** F
MAP-WRONG	.39	.46	1.68
MRT-RIGHT	.07	-1.39	-1.67
MRT-WRONG	1.02	1.07	.15
SAT-MATH	.52	.52	.06
SAT-VERBAL	.57	.13	1.42
GPA	3.26*** F M	4.44*** F	5.24****f M
GENDER	-.49	4.38*	-.99
PRIOR	3.28*	-2.71**	2.56
MAJOR	3.42* F	2.15	5.54** M
R^2 =	[.2449]	[.2562]	[.3400]
Exam 2 =	82.13****F M	117.31****F M	67851****F M
MAP-RIGHT	.83	1.01	2.36 f
MAP-RIGHT	1.56	.18	.83
MRT-RIGHT	-.23	-4.02*	-1.67
MRT-WRONG	.56	1.30	-.22
SAT-MATH	-.55	-1.57	.56
SAT-VERBAL	-1.43	-.24	2.49*
GPA	4.52****F M	8.43****f M	5.83****F m
GENDER	-.15	7.25*	-4.53
PRIOR	2.19	-3.46	1.20
MAJOR	4.40** F	5.65	1.67
R^2 =	[.1993]	[.2432]	[.3255]

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

[18]Females: F = sig. (p < .05) f = sig. (p < .10)
Males: M = sig. (p < .05) m = sig. (p < .10)

Exam 3 =	83.58****	90.32****F M
MAP-RIGHT	1.09	.82
MAP-WRONG	.08	1.20
MRT-RIGHT	-1.90*	-.66
MRT-WRONG	.22	.06
SAT-MATH	.10	.59
SAT-VERBAL	-.93	.19
GPA	4.48****F	3.03*** F M
GENDER	2.93	.22
PRIOR	-2.72	.70
MAJOR	2.91	2.09
R^2 =	[.2935]	[.2484]
<u>Exam 4 =</u>	111.10****F M	109.54****F M
MAP-RIGHT	-.72	1.96
MAP-WRONG	.76	2.52
MRT-RIGHT	.92	.15
MRT-WRONG	3.19*	-.35
SAT-MATH	3.14	.88
SAT-VERBAL	-.22	1.96
GPA	8.74****F M	8.85****F M
GENDER	.92	1.29
PRIOR	3.39	-.04
MAJOR	5.58	2.90
R^2 =	[.3217]	[.3500]

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

Principles I but is unrelated to Principles II and Intermediate. MAJOR is not significant for any course.

<u>Intermediate--full models with MAP and MRT percentages</u>. Table 34 reports models which replace VSA components with VSA percentages. Results are comparable. Neither MAP-% nor MRT-% are significant for the two-gender sample, although MRT-% is significant for males at p < .10.

Table 34
Results of regression full models
of Principles I, II, and Intermediate performance scores,
on VSA percentage scores, tertiary sample[19]
N = 94 (53 females, 41 males)

PANEL A: Wilks' Lambda p-values (p < .05 underlined):

	Prin I	Prin II	Inter
MAP-%	.5713	.6546	.7515
MRT-%	.5511	.2212 m	.4478
SAT-MATH	.3562	.4953	.6604
SAT-VERBAL	.2318	.7019	.2346
GPA	.0001 F M	.0001 F M	.0003 F M
GENDER	.6246	.3378	.1259
PRIOR	.0304 F	.5203	.6994
MAJOR	.3410 F	.7338	.1170

PANEL B: Univariate parameter estimates:

	Prin I	Prin II	Inter
Exam 1 =	85.46****F M	78.99****F M	74.60****F M
MAP-%	.12	-.17	.83
MRT-%	-1.23	-2.02*	-1.31
SAT-MATH	.65	.68	1.02
SAT-VERBAL	.68	.00	1.47 f
GPA	3.38*** F M	4.69****F m	4.97*** F M
GENDER	-.05	4.00*	.07
PRIOR	3.26* F	-2.87	1.99
MAJOR	3.34* F	2.01	4.79* m
R^2 =	[.2427]	[.2545]	[.2982]
Exam 2 =	82.04****F M	118.24****F M	67.95****F M
MAP-%	-.71	1.24	.57
MRT-%	-.99	-4.33** M	-.01
SAT-MATH	-.32	-.94	1.15
SAT-VERBAL	-1.34	-.77	2.48*
GPA	4.57****F M	9.02****f M	5.35*** F
GENDER	.37	6.20*	-4.39*
PRIOR	2.09	-4.02	1.21
MAJOR	4.14* F	5.03 m	1.18
R^2 =	[.1913]	[.2374]	[.3069]

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

[19]Females: F = sig. (p < .05) f = sig.(p < .10)
Males: M = sig. (p < .05) m = sig. (p < .10)

PANEL B, continued: Univariate parameter estimates

Exam 3 =	83.93****F M	90.36****F M
MAP-%	.74	-.49
MRT-%	-1.50	-.46
SAT-MATH	.53	.83
SAT-VERBAL	-1.15	.19
GPA	4.59****F M	2.96*** F M
GENDER	2.65	.43
PRIOR	-2.99	.60
MAJOR	2.60	1.83
R^2 =	[.2781]	[.2341]
Exam 4 =	110.82****F M	109.23****F M
MAP-%	-.93	-.79
MRT-%	-1.37	.03
SAT-MATH	2.66	1.35
SAT-VERBAL	.26	2.16
GPA	8.60****F M	8.63****F M
GENDER	.70	2.55
PRIOR	4.14	-.17
MAJOR	5.64 M	2.38
R^2 =	[.2993]	[.3385]

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

<u>Analysis for subjects taking Cost</u>. Tables 35, 36, and 37 compare parameter estimates for Principles I, II and Cost models using the reduced sample of subjects who continued to Cost (N = 75). Table 35 reports models for which VSA scores explain some of the variance in the first two EXAMS of these three courses. Tables 36 (VSA components) and 37 (VSA percentages) compare results of full models to explain grade variance.

<u>Cost--partial models</u>. Table 35 reports results of regressions for three courses for MAP and MRT component

Table 35
Results of regression partial models
of Principles I, II and Intermediate performance scores, on VSA component or percentage scores, tertiary sample[20]

N = 75 (40 females, 35 males)

PANEL A:

Wilks' Lambda p-values (p < .05 underlined):

	Prin I	Prin II	Cost
MAP-RIGHT	.9455	.9958	.4704
MAP-WRONG	.4729	.9152	.4802
MRT-RIGHT	.9003	.3373	.3180
MRT-WRONG	.2652	.7993	.5043

Univariate parameter estimates:

	Prin I	Prin II	Cost
Exam 1 =	88.47****F M	80.13****F M	73.85****F M
MAP-RIGHT	.14	.12	-1.02
MAP-WRONG	.04	.72	-2.46
MRT-RIGHT	.76	-1.13	.52
MRT-WRONG	.94	-.09	-2.01
R^2 =	[.0163]	[.0131]	[.0480]
Exam 2 =	85.11****F M	119.85****F M	75.69***F M
MAP-RIGHT	-.29	.50	-2.71
MAP-RIGHT	1.05	.01	-1.04
MRT-RIGHT	.15	-2.00	2.69 F
MRT-WRONG	.26	-.50	-1.24
R^2 =	[.0211]	[.0090]	[.0545]
Exam 3 =	85.65****F M	91.73****F M	
MAP-RIGHT	-.88	.30	
MAP-WRONG	-1.15	.70	
MRT-RIGHT	-.27	.27	
MRT-WRONG	-1.28	-.60	
R^2 =	[.0384]	[.0180]	
Exam 4 =	116.95****F M	112.92****F M	
MAP-RIGHT	-1.13	-.10	
MAP-RIGHT	.90	1.35	
MRT-RIGHT	1.11	2.68	
MRT-WRONG	1.17	-2.47	
R^2 =	[.0158]	[.0528]	

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

[20]Females: F = sig. (p < .05) f = sig.(p < .10)
Males: M = sig. (p < .05) m = sig. (p < .10)

PANEL B:

Wilks' Lambda p-values:

	Prin I	Prin II	Cost
MAP-%	.4606	.7575	.2010
MRT-%	.4496 F	.0474	.0308 F

	Prin I	Prin II	Cost
Exam 1 =	88.47****F M	80.13****F M	73.85****F M
MAP-%	.05	-.59	1.94
MRT-%	-.30	-.29	3.10*
R^2 =	[.0012]	[.0039]	[.0728]
Exam 2 =	85.11****F M	119.85****F M	75.69****F M
MAP-%	-1.01	.79	-.69
MRT-%	-.04	-.18	4.21** F
R^2 =	[.0136]	[.0018]	[.0901]
Exam 3 =	85.65****F M	91.73****F M	
MAP-%	.58	-.38	
MRT-%	1.28 F	1.27	
R^2 =	[.0275]	[.0346]	
Exam 4 =	116.95****F M	112.92****F M	
MAP-%	-1.54	-1.09	
MRT-%	1.47	4.99** M	
R^2 =	[.0156]	[.0732]	

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

scores in Panel A. No VSA measure is significant. Panel B reports partial models using percentage scores. MAP-% is unrelated to all courses. MRT-% is not related to Principles I, but is related to Principles II and Cost.

Cost--full models with MAP and MRT component scores. Table 36 reports results of regressions for Principles I, II and Cost with the full vector of independent variables including VSA components. VSA measures are not significant for any course at conventional levels. SAT-MATH is related to Cost performance. GPA is related to all courses.

Table 36
Results of regression full models of Principles I, II and Cost performance scores on VSA component scores, tertiary sample[21]

N = 75 (40 females, 35 males)

PANEL A: Wilks' Lambda p-values:

	Prin I	Prin II	Cost
MAP-RIGHT	.7178	.9770	.5695
MAP-WRONG	.1517	.9458	.2404
MRT-RIGHT	.9217	.2313	.5462
MRT-WRONG	.2424	.6104	.8653
SAT-MATH	.3771	.3268	.0358 m
SAT-VERBAL	.7200	.8877	.3115
GPA	.0002 F	.0001 F M	.0022 F
GENDER	.3361	.5581	.5856
PRIOR	.3907	.4195	.8921
MAJOR	.7535 F	.5684	.3090

PANEL B: Univariate parameter estimates:

	Prin I	Prin II	Cost
Exam 1 =	87.30****F M	80.49****F M	72.87****F M
MAP-RIGHT	.44	.09	-1.97
MAP-WRONG	.44	.85	-2.94 f
MRT-RIGHT	.50	-1.67	.75
MRT-WRONG	1.78	2.18	.81
SAT-MATH	1.14	.92	4.65** M
SAT-VERBAL	.64	.08	.22
GPA	3.38** F	4.49*** F	5.49** F M
GENDER	-.83	2.77	.60
PRIOR	1.64	-4.21	1.37
MAJOR	2.07 F	1.20	.08
R^2 =	[.2485]	[.2834]	[.3445]
Exam 2 =	84.79****F M	117.25****F M	74.61****F M
MAP-RIGHT	.99	1.47	-1.92
MAP-WRONG	2.04	1.14	-.01
MRT-RIGHT	-.23	-3.66	1.95 f
MRT-WRONG	.70	3.33	-.04
SAT-MATH	-.75	-.85	1.96
SAT-VERBAL	-.76	-.98	2.61 F
GPA	4.24****F M	11.86****F M	4.57** F
GENDER	-2.07	6.48	-2.66
PRIOR	.73	-5.57	1.32
MAJOR	2.57	6.43	4.66
R^2 =	[.2376]	[.4055]	[.3414]

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

[21]Females: F = sig. (p < .05) f = sig.(p < .10)
Males: M = sig. (p < .05) m = sig. (p < .10)

Exam 3 =	85.82****F M	91.13****F M
MAP-RIGHT	-.94	.29
MAP-WRONG	-1.13	.79
MRT-RIGHT	-.55	.20
MRT-WRONG	.48	.80
SAT-MATH	.69	1.38
SAT-VERBAL	-.55 F	-.46
GPA	3.73*** F	3.81****F m
GENDER	2.14	.35
PRIOR	-2.74	.24
MAJOR	.50	.90
R^2 =	[.2630]	[.3363]
<u>Exam 4 =</u>	116.79****F M	111.24****F M
MAP-RIGHT	.12	.15
MAP-RIGHT	2.02	1.80
MRT-RIGHT	.40	2.03
MRT-WRONG	3.91*	1.32
SAT-MATH	1.66	2.94
SAT-VERBAL	.72 F	.78
GPA	9.42****F M	10.89****F M
GENDER	-1.57	1.93
PRIOR	-.32	.07
MAJOR	2.91	2.07
R^2 =	[.3921]	[.4536]

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

<u>Cost--full models with VSA percentages</u>. Table 37 reports regressions for three courses using the full independent variable vector including MAP-% and MRT-%. Neither percentage measure is related to Principles I, but MRT-% is related to Principles II. Results for the non-VSA variables are comparable to those reported in the previous table for the VSA component model. Again, GPA is explanatory for all courses.

In models for these tertiary samples, the VSA-performance relationships reported for Principles I and Principles II, for the primary and secondary samples, are no longer seen.

Table 37
Results of regression full models
of Principles I, II, and Cost performance scores
on VSA percentage scores, tertiary sample[22]

---------------N = 75 (40 females, 35 males)---------------

PANEL A:
Wilks' Lambda p-values (p < .05 underlined):

	Prin I		Prin II		Cost	
MAP-%	.3957		.8102		.1238	
MRT-%	.6414		.0322	m	.2859	
SAT-MATH	.5748		.3967		.0488	M
SAT-VERBAL	.5101		.8254		.1836	f
GPA	.0002		.0001	F M	.0030	F
GENDER	.6858		.5440		.5379	
PRIOR	.3514		.2706		.8151	
MAJOR	.7356		.6014		.2423	

PANEL B: Univariate parameter estimates:

	Prin I		Prin II		Cost	
Exam 1 =	87.01****	F M	80.92****	F M	73.25****	F M
MAP-%	-.20		-.75		1.84	
MRT-%	-1.63		-3.25**		.28	
SAT-MATH	1.04		.94		4.03**	M
SAT-VERBAL	.85		.04		.36	
GPA	3.44**	F	4.86***	F	5.49**	F m
GENDER	-.44		2.19		-.58	
PRIOR	1.66		-4.47*		1.49	
MAJOR	2.35	F	1.09		.40	
R^2 =	[.2389]		[.2891]		[.3368]	
Exam 2 =	84.56****	F M	117.99****	F M	74.35****	F M
MAP-%	-1.16		.27		-1.26	
MRT-%	-1.18		-5.90**	m	2.45	
SAT-MATH	-.54		-.37		1.19	
SAT-VERBAL	-.69		-1.26		3.05*	F
GPA	4.29***	F M	12.51****	F M	4.15**	F
GENDER	-1.42		5.84		-3.13	
PRIOR	.68		-6.05		1.99	
MAJOR	2.43		5.87		5.06	
R^2 =	[.2275]		[.4125]		[.3478]	

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

Panel B continues next page.

[22]Females: F = sig. (p < .05) f = sig.(p < .10)
Males: M = sig. (p < .05) m = sig. (p < .10)

PANEL B, continued: Univariate parameter estimates:

	Prin I	Prin II
Exam 3 =	86.30****F M	90.92****F M
MAP-%	.52	-.47
MRT-%	-.81	.32
SAT-MATH	.53	1.30
SAT-VERBAL	-.65 F	-.28
GPA	3.95*** F	3.70****F m
GENDER	1.34	.52
PRIOR	-2.98	.52
MAJOR	.53	.92
R^2 =	[.2512]	[.3219]
Exam 4 =	116.25****	110.45****F M
MAP-%	-1.91	-1.46
MRT-%	-2.11	.29
SAT-MATH	1.10	2.58
SAT-VERBAL	1.40 F	1.32
GPA	9.15****F M	10.56****F M
GENDER	-1.70	2.80
PRIOR	.63	.57
MAJOR	3.26	2.45
R^2 =	[.3612]	[.4377]

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

Those subjects continuing to Cost showed a relationship of MRT-% to Cost EXAMS, while those continuing to Intermediate showed no relationship of VSA to their EXAMS. This may be attributable to sample differences, but the coefficients for Intermediate and Cost cannot be compared since they arise from nonidentical samples. The next section addresses this issue.

Comparison across four courses of VSA and performance

To compare the role of aptitudes in Intermediate and Cost, separate regressions were run for all four courses (Principles I and II, Intermediate and Cost) using a third tertiary sample--those 71 subjects who took both Intermediate and Cost. Tables 38 and 39 provide Wilks' Lambda p-values and univariate parameter estimates for the EXAMS taken in each course. Regressions include all variables included in previous regressions. Table values are presented side by side for course comparisons. The VSA component partial model had neither multivariate nor univariate significance and is not included in the tables. MAP and MRT component measures. The VSA component full model is presented in Table 38. With the small sample (N = 71), the only significant VSA measure is MAP-RIGHT, which is related to Intermediate only. GPA is related to all courses.

Table 38[23]
Results of regression full models
of Principles I, II, Intermediate, Cost performance scores
on VSA component scores, tertiary sample[24]

N = 71

PANEL A:

Wilks' Lambda p-values (p < .05 underlined)

	Prin I	Prin II	Inter	Cost
MAP-RIGHT	.7738	.8198	.0478	.8293
MAP-WRONG	.1841	.8771	.1888	.4315
MRT-RIGHT	.9078	.2118	.4895	.7858
MRT-WRONG	.2796	.5841	.7454	.8461
SAT-MATH	.7118	.3279	.6628	.0803
SAT-VERBAL	.8352	.7414	.5622	.2459
GPA	.0003 F	.0001 FM	.0001 FM	.0010 F
GENDER	.3587	.6033	.2084	.6080
PRIOR	.4790	.3588	.9068	.8218
MAJOR	.6459	.5367	.1103 m	.3601

Univariate parameter estimates

	Prin I	Prin II	Inter	Cost
EXAM 1 =	87.56****FM	80.32****FM	77.05****FM	72.91****FM
MAP-RIGHT	1.02	.89 F	3.96** f	-.93
MAP-WRONG	.86	1.54 F	2.61	-2.19**
MRT-RIGHT	.27	-2.13	-1.64	.42
MRT-WRONG	1.66	2.38* f	.91	.92
SAT-MATH	.32	.51	.10	3.99**
SAT-VERBAL	.17	-.39	.03	-.67
GPA	3.68** F	5.30*** FM	6.47****FM	6.29***FM
GENDER	-1.28	3.11	-2.60	.65
PRIOR	1.78	-4.14	.92	1.09
MAJOR	2.55 F	1.64	5.71** M	.92
R^2 =	[.2410]	[.2914]	[.3757]	[.3361]

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

Panel B continued next page.

[23]VSA component partial model shows neither multivariate nor univariate significance and is therefore not included in table.

[24]Females: F = sig. (p < .05) f = sig.(p < .10)
Males: M = sig. (p < .05) m = sig. (p < .10)

Table 38, Panel B continued

PANEL B:	Univariate parameter estimates			
	Prin I	Prin II	Inter	Cost
EXAM 2 =	84.81****FM	118.82****FM	70.69****FM	74.33****FM
MAP-RIGHT	1.25	3.29 f	.89	-1.33
MAP-WRONG	2.41*	2.23	.51	.46
MRT-RIGHT	-.35	-3.88*	-1.71	1.25
MRT-WRONG	.81	3.24	.13	-.02
SAT-MATH	-.77	-1.85	1.26	1.00
SAT-VERBAL	-1.08	-2.47	1.50	2.58 f
GPA	4.57*** FM	12.02****FM	5.61** F	5.25*** Fm
GENDER	-1.96	6.06	-5.40*	-2.61
PRIOR	.64	-6.44	-.03	2.15
MAJOR	3.14	6.90	3.42	4.78
R^2 =	[.2429]	[.3831]	[.3571]	[.3303]
EXAM 3 =	85.71****FM	91.26****FM		
MAP-RIGHT	-.52	.51		
MAP-WRONG	-.78	1.01		
MRT-RIGHT	-.86	-.02		
MRT-WRONG	.55	.82		
SAT-MATH	.32	1.01		
SAT-VERBAL	-.76	-.53		
GPA	4.16*** F	3.98****Fm		
GENDER	2.24	.33		
PRIOR	-2.53	.48		
MAJOR	.80	1.07		
R^2 =	[.2615]	[.3090]		
EXAM 4 =	117.39****	111.81****		
MAP-RIGHT	.99	.23		
MAP-WRONG	2.78 F	2.06		
MRT-RIGHT	.17	1.81		
MRT-WRONG	4.05*	1.39		
SAT-MATH	1.21	2.58		
SAT-VERBAL	-.12	.67		
GPA	9.84****FM	10.86****FM		
GENDER	-1.57	1.91		
PRIOR	-.65	.37		
MAJOR	3.75	2.64		
R^2 =	[.3720]	[.4226]		

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

VSA percentage measures. Table 39 presents results of VSA percentage partial (Panel A) and full (Panel B) models for all four courses. In the partial model, neither MAP-% nor MRT-% is related to the Principles I or Principles II vectors at conventional levels, although the relationship of MRT-% and Principles II is close at $p < .0600$. Neither VSA percentage measure is related to the Intermediate vector, but MRT-% is related to Cost.

In the full models, MRT-% is related to performance in Principles II, but to no other course. Neither the Intermediate nor Cost grade vector is related to VSA. GPA is related to all courses.

Table 39
Results of regression partial and full models of Principles I, II, Intermediate, Cost performance scores on VSA percentage scores, tertiary sample[25]

N = 71

PANEL A:	Wilks' Lambda p-values (p < .05 underlined)			
	Prin I	Prin II	Inter	Cost
MAP-%	.4447	.6408	.7057	.2055
MRT-%	.4998 F	.0600	.2688	.0363

	Univariate parameter estimates			
	Prin I	Prin II	Intermediate	Cost
EXAM 1 =	88.76****FM	80.37****FM	78.35****FM	74.10****FM
MAP-%	.21	-.48	1.14	2.15
MRT-%	-.33	-.35	.86	3.30** F
R^2 =	[.0022]	[.0031]	[.0161]	[.0858]
EXAM 2 =	85.34****FM	121.08****FM	69.34****FM	75.89****FM
MAP-%	-.96	1.41	.79	-.53
MRT-%	.07	.09	2.38	4.04** F
R^2 =	[.0117]	[.0061]	[.0364]	[.0847]
EXAM 3 =	85.83****	92.04****		
MAP-%	.70	-.24		
MRT-%	1.24 F	1.20		
R^2 =	[.0276]	[.0301]		
EXAM 4 =	117.69****	113.87****		
MAP-%	-1.26	-.71		
MRT-%	1.77	5.06** M		
R^2 =	[.0165]	[.0742]		

PANEL B:	Wilks' Lambda p-values (p < .05 underlined)			
	Prin I	Prin II	Intermediate	Cost
MAP-%	.3908	.7068	.9397	.1593
MRT-%	.6633	.0454	.4098	.4155
SAT-MATH	.8701	.4525	.6296	.0757 M
SAT-VERB	.6301	.6651	.6657	.1650
GPA	.0004 FM	.0001 FM	.0005 FM	.0014 F
GENDER	.6736	.5486	.0820	.5518
PRIOR	.3806	.2489	.7847	.7038
MAJOR	.6637 F	.6001	.1895	.3035

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

[25]Females: F = sig. (p < .05) f = sig. (p < .10)
Males: M = sig. (p < .05) m = sig. (p < .10)

Table 37, Panel C: Univariate parameter estimates

	Prin I	Prin II	Inter	Cost
EXAM 1 =	87.15****FM	80.67****FM	76.42****FM	73.05****FM
MAP-%	-.20	-.91	.42	1.72
MRT-%	-1.58	-3.56** f	-1.49	.21
SAT-MATH	.32	.72	.91	3.58** M
SAT-VERBAL	.55	-.35	.36	-.57
GPA	3.63** F	5.50*** F	5.90****FM	6.16*** Fm
GENDER	-.78	2.68	-1.55	-.13
PRIOR	2.06	-4.24	1.69	1.42 m
MAJOR	2.65 F	1.37	5.01* M	1.11
R^2 =	[.2245]	[.2911]	[.3233]	[.3344]
EXAM 2 =	84.54****FM	119.13****FM	70.90****FM	74.12****FM
MAP-%	-1.32	.39	.22	-1.33
MRT-%	-1.31	-5.54**	-.06	2.12**
SAT-MATH	-.54	-1.03	1.51	.43
SAT-VERBAL	-.85	-2.40	1.42	2.87 F
GPA	4.53*** FM	12.09****FM	5.27** F	4.83** F
GENDER	-1.31	5.97	-5.93**	-3.01
PRIOR	.74	-6.36	.38	2.79
MAJOR	2.90	6.08	2.98	5.01
R^2 =	[.2290]	[.3766]	[.3445]	[.3400]
EXAM 3 =	86.14****	91.05****		
MAP-%	.44	-.50		
MRT-%	-1.00	.41		
SAT-MATH	.29	.98		
SAT-VERBAL	-.93	-.30		
GPA	4.36*** F	3.79****FM		
GENDER	1.59	.45		
PRIOR	-2.77	.83		
MAJOR	.76	1.00		
R^2 =	[.2516]	[.2923]		
EXAM 4 =	116.72****	111.05****		
MAP-%	-2.07	-1.62		
MRT-%	-2.11	.10		
SAT-MATH	.76	2.20		
SAT-VERBAL	.82	1.34		
GPA	9.26****FM	10.54****FM		
GENDER	-1.61	2.64		
PRIOR	.66	1.03		
MAJOR	3.89	2.86		
R^2 =	[.3334]	[.4069]		

[*** p = .0001, *** p = .001, ** p < .05, * p < .10]

Four course regressions--summary. In partial regressions, the only VSA-performance relationship is MRT-% and Cost EXAMS. In full regressions, the only VSA-performance relationships are MRT-% and Principles II and MAP-RIGHT and Intermediate.

Course differences in the aptitude-performance relationship--summary. In analyses with the secondary sample, MAP-RIGHT was related to Principles II but not to Principles I. [MRT-RIGHT and MRT-% were related to both courses.] With the smaller tertiary sample, however, tests are no longer powerful enough to identify some of the VSA effects found previously in Principles I and II. Even with this small sample size (N = 71), however, partial and full regressions have reported course differences in the VSA-performance relationship.

Gender differences in the course differences. In separate regressions by gender, gender differences also appear. For the secondary sample, females showed VSA-performance relationships for Principles I, while males showed such relationships for Principles II. In the tertiary sample, females have the only VSA-performance relationships observable at this small sample size (35 males, 40 females). For females only, MRT-% is related to Principles I in partial regressions. Although gender differences were only hypothesized for the VSA-performance relationship in Principles I, it is clear that such differences exist in Principles II, also.

Chapter 6 -- Discussion and Conclusions

This research has examined the role of VSA in undergraduate student behavior and performance in accounting coursework. This chapter provides a brief summary of results of tests of hypotheses and provides discussion and conclusions related to these findings. Summary tables are provided to supplement the text.

VSA and completing and continuing accounting coursework. Table 40 summarizes significant tests of the hypothesized relationship between VSA and completing Principles I. Results show that subjects completing Principles I have higher MAP-RIGHT and MRT-RIGHT scores than do subjects dropping the course.[26] Results also show that subjects in the bottom thirds of the MAP-RIGHT and/or MRT-RIGHT sample distributions are less likely to complete the course than would be expected, while subjects in the top thirds of the MAP-RIGHT and/or

---

[26] To examine the consistency of VSA's relationship to dropping behavior with VSA's relationship to performance on the first exam (which some dropping students had taken before dropping), separate regressions were run for each VSA variable with EXAM 1 score. For MAP variables, at least, the relationship to completion behavior, identified in *t*-tests and Chi-square, is consistent with the relationship to the EXAM 1 score. MAP-RIGHT was related to the first exam score, while MAP-WRONG and MAP-% were not.

For MRT variables, however, no such consistency is found. Whereas completion was related to MRT-RIGHT in *t*-tests and Chi-square, univariate regressions showed that MRT-WRONG and MRT-% were related to performance on EXAM 1. In multivariate partial regressions, however, MRT-RIGHT is related to Principles I performance in general.

Table 40
Summary of tests of a relationship between VSA and completing Principles I

Panel A: *t*-tests of aptitudes across completion status

	No difference	Completers higher
MAP-RIGHT		x
MAP-WRONG	x	
MAP-%	x	
MRT-RIGHT		x
MRT-WRONG	x	
MRT-%	x	
SAT-MATH		x
SAT-VERBAL		x
GPA		x

Panel B: Chi-square tests of independence

	Actual completion rate		
	< Expected	= Expected	> Expected
MAP-RIGHT			
Low scorers	x		
Medium scorers	x		
High scorers			x
MRT-RIGHT			
Low scorers	x		
Medium scorers		x	
High scorers			x

MRT-RIGHT distributions are more likely to complete the course than would be expected. These findings support the hypothesized relationship between completing behavior and VSA. The significance of MAP-RIGHT and MRT-RIGHT but not MAP-% and MRT-% suggests that completing behavior is related to VSA processing speed but not to VSA processing accuracy.

Table 41 summarizes results for tests of the relationship between VSA and continuing to Principles II. Differences between continuers and noncontinuers are significant for MRT-% at $p < .05$ and for MRT-WRONG at $p < .0512$. This finding is not surprising; high performers in Principles I and continuers to Principles II have higher MRT-% and lower MRT-WRONG than do low performers and noncontinuers. Consistent with this is the difference between continuers and noncontinuers on Principles I total score.

Results also indicate that continuing to Principles II is not independent of VSA levels, as measured by MAP-WRONG and MRT-WRONG.[27] Subjects missing the most items on the MAP and/or MRT tests, or high scorers, continue to Principles II at a lower rate than would be anticipated. Subjects missing few items on the MRT continue to Principles II at a higher rate than expected. The unusual finding, however, is that subjects missing few items on the MAP continue to Principles II at a <u>lower</u> rate than expected. For MAP-WRONG scores, low and high scorers continue less frequently and medium scorers continue more frequently than expected. There is no explanation for this anomaly.

---

[27]The significance of MRT-WRONG is evident in Chi-square results discussed here and *t*-test results discussed in the previous paragraph. Significance for MRT-% is found only in *t*-tests because Chi-square is not an appropriate test for percentage measures.

Table 41
Summary of results of tests of a relationship between VSA and continuing to Principles II

Panel A: *t*-tests of aptitudes across continuance status

	Continuers lower	No difference	Continuers higher
MAP-RIGHT		x	
MAP-WRONG		x	
MAP-%		x	
MRT-RIGHT		x	
MRT-WRONG	$p < .0512$		
MRT-%			x
SAT-MATH			x
SAT-VERBAL		x	
GPA			x
Prin I TOTAL			x

Panel B: Chi-square tests of independence

	Actual continuance rate		
	< Expected	= Expected	> Expected
MAP-WRONG			
Low scorers	x		
Medium scorers			x
High scorers	x		
MRT-RIGHT			
Low scorers		$p < .053$	
Medium scorers	$p < .053$		
High scorers			$p < .053$
MRT-WRONG			
Low scorers			x
Medium scorers		x	
High scorers	x		

VSA and performance. Table 42 summarizes results for the tests of the relationship between VSA and performance in Principles I. Results indicate multiple significant tests of a VSA-performance relationship in accounting coursework.

For the sample of those subjects enrolled in Principles I, both categorical and continuous models generally indicate significance for two VSA measures--MRT-RIGHT and MRT-%. MAP-%, however, is significant in categorical models but not in continuous models.[28]

In regressions, some Principles I performance score variance is explained by errors and accuracy in spatial relations, as measured by MRT.[29] All MRT measures are related to Principles I performance in partial regression models, while MRT-WRONG and MRT-% are significant in full models. Performance in Principles I is related negatively to the number of MRT errors and positively to accuracy on the MRT. The grade variance explained previously by MRT-RIGHT, in partial models, is also explainable by one or more of the covariates included in full models. The significance of MRT-WRONG and MRT-%, in models which include other

---

[28]MRT-WRONG is significant in continuous models but is not included in categorical models.

[29]Spatial scanning, as measured by the MAP variables, is not related to performance in any Principles I regression model. MAP's relationship to dropping (N = 576) and nonrelationship to performance (N = 513) may be a function of underlying sample differences.

Table 42
Summary of tests of a relationship
between VSA and performance in Principles I

Partial MANOVA models:

Prin I scores =	MAP-RIGHT	$\underline{\text{MRT-RIGHT}}$ (e,h)	MAP-RIGHT*MRT-RIGHT	
Prin I scores =	$\underline{\text{MAP-\%}}$ (e)	$\underline{\text{MRT-\%}}$ (e,h)	MAP% * MRT-%	

Full MANCOVA models:

Prin I scores =	MAP-RIGHT	MRT-RIGHT	
	$\underline{\text{SAT-MATH}}$	$\underline{\text{SAT-VERBAL}}$	$\underline{\text{GPA}}$
	$\underline{\text{GENDER}}$	$\underline{\text{PRIOR}}$	$\underline{\text{MAJOR}}$
Prin I scores =	MAP-%	MRT-%	
	$\underline{\text{SAT-MATH}}$	$\underline{\text{SAT-VERBAL}}$	$\underline{\text{GPA}}$
	$\underline{\text{GENDER}}$	$\underline{\text{PRIOR}}$	$\underline{\text{MAJOR}}$

Partial REGRESSION models:

Prin I scores =	MAP-RIGHT	MAP-WRONG	$\underline{\text{MRT-RIGHT}}$ (h)	$\underline{\text{MRT-WRONG}}$ (e)
Prin I scores =	MAP-%	$\underline{\text{MRT-\%}}$ (e,h)		

Full REGRESSION models:

Prin I scores =	MAP-RIGHT	MAP-WRONG	MRT-RIGHT	$\underline{\text{MRT-WRONG}}$ (h)
	$\underline{\text{SAT-MATH}}$	$\underline{\text{SAT-VERBAL}}$	$\underline{\text{GPA}}$	
	$\underline{\text{GENDER}}$	$\underline{\text{PRIOR}}$	$\underline{\text{MAJOR}}$	
Prin I scores =	MAP-%	$\underline{\text{MRT-\%}}$ (e,h)		
	$\underline{\text{SAT-MATH}}$	$\underline{\text{SAT-VERBAL}}$	$\underline{\text{GPA}}$	
	$\underline{\text{GENDER}}$	$\underline{\text{PRIOR}}$	$\underline{\text{MAJOR}}$	

Legend:
Underlined variables indicate multivariate significance.
(e) indicates univariate significance for EXAMS.
(h) indicates univariate significance for HW/QUIZ.

aptitude covariates, supports the hypothesized relationship of VSA and performance in Principles I.

The practical significance of the VSA-relationship, however, appears small. R^2 for partial models was generally less than 5%. It is impossible to say at this point whether the relatively small R^2 associated with the VSA components and percentages is an indication of the small practical value of VSA in accounting education, or if it is more attributable to the particular instructional methods and course requirements of this institution. The courses at this institution may be designed in such a way that they often obviate the need for VSA by avoiding certain activities for which VSA would be helpful. It would be unwarranted to conclude, based on small R^2, that VSA is of no practical importance in accounting education. Future research might be able to demonstrate that alternate course designs increase the moderating effect of student VSA. That is, some in-class and outside-class experiences might better make use of the VSA that students have.

Gender differences in students enrolled in Principles I. As suggested by the literature, males score higher on three of the four VSA measures indicating skill at VSA tasks--MAP-RIGHT, MRT-RIGHT, and MRT-%. Female scores are higher than males on MRT-WRONG. Females outscore males on EXAM 2, HW/QUIZ and COMPUTER, while males outscore females on EXAMS 3 and 4.

There is no gender difference in total points earned in Principles I. In this sample, however, males scores are higher than female scores on SAT-MATH and SAT-VERBAL. It is not known if this is typical of Principles I students at other institutions.

Table 43 summarizes results of tests of the hypothesized gender difference in the VSA relationship to Principles I performance. In partial regressions, females and males both exhibit relationships of performance to MRT-WRONG and MRT-%. However, these relationships for females are for EXAMS and HW/QUIZ, whereas the only VSA relationships for males are for EXAMS. In full regressions, only females still exhibit VSA-performance relationships in the presence of covariates;

Table 43
Summary of tests of a gender difference in the relationship between VSA and Principles I performance

	Significant VSA-performance relationships for Females	Males
Partial regression:	MRT-WRONG (HW/Quiz)	MRT-WRONG (Exams)
	MRT-% (Exams, HW/Quiz)	MRT-% (Exams)
Full regression:	All covariates MRT-WRONG (HW/Quiz) MRT-% (HW/Quiz, Computer)	All covariates except MAJOR

MRT-WRONG is related to HW/QUIZ, while MRT-% is related to HW/QUIZ and COMPUTER.[30] For males, the grade variance explained by VSA in partial models is also explainable by one or more of the covariates included in full models. This finding supports the hypothesized gender difference in the relationship between VSA and Principles I performance.

<u>The relationship of VSA to performance across courses</u>. Table 44 summarizes results of tests of course differences in the VSA-performance relationship. For those students who continued to Principles II, partial regression results indicate that MRT-RIGHT and MRT-% are related to performance in both courses while MAP-RIGHT is related only to performance in Principles II. In univariate analysis, MRT-% is related to EXAMS in both courses and to HW/QUIZ only in Principles I. MRT-RIGHT is related to HW/QUIZ in Principles I and is related only to the grade vector in Principles II. Whereas for the primary sample, MRT-WRONG is related to Principles I scores, MRT-WRONG is not related to Principles I performance in the secondary sample.[31] This is more understandable when it is

---

[30]It is particularly interesting that full regressions show significance for females for two of the three performance measures on which female scores are higher than are male scores--HW/QUIZ and COMPUTER. The third performance measure for which females score higher than males--EXAM 2--shows a marginally significant relationship to MRT-WRONG at $p < .0682$ for females, while the p-value for males is .5386.

[31]For Principles II, secondary sample, partial model multivariate significance for MRT-WRONG is only $p < .1084$, even though MRT-WRONG is negatively related to EXAMS 1, 3 and 4 at

recalled that MRT-WRONG was a variable on which differences were found between continuers and noncontinuers. The different findings across samples for MRT-WRONG and Principles I scores may be attributable to sample differences.

Table 44
Summary of regression tests of course differences in the VSA-performance relationship

Secondary sample (N = 342):

	Significant regression relationships	
	Partial Regression	Full Regression
Prin I	MRT-RIGHT (hw/quiz)	--
	MRT-% (exams, hw/quiz)	
Prin II	MAP-RIGHT (exams)	--
	MRT-RIGHT[32]	
	MRT-% (exams)	

Tertiary sample (N = 71):

	Significant regression relationships	
	Partial Regression	Full Regression
Prin I	--	--
Prin II	--	MRT-% (exams)
Interm.	--	MAP-RIGHT (exams)
Cost	MRT-% (exams)	--

$p < .0204$, $.0046$, and $.0208$.

[32]Although MRT-RIGHT is multivariately significant at $p < .0476$, it is not univariately related to any Principles II performance score. Instead, MRT-RIGHT is related to the vector of performance scores.

For this secondary sample, no VSA measures are significant in full regressions. For this sample, at least, the grade variance explained by VSA measures in partial models is explainable by one or more covariates included in full models.

When examining results for those subjects who continued to both Intermediate and Cost, it should be recalled that only EXAMS are examined for relationships to VSA in this sample. Accordingly, the significant relationships found previously for MRT-RIGHT and MRT-% (to Principles I HW/QUIZ) are no longer found. VSA is therefore unrelated to performance in Principles I, but is related to EXAM scores in Principles II, Intermediate and Cost. MRT-% is related to Principles II (in full regressions) and Cost (in partial regressions), while MAP-RIGHT is related to performance in Intermediate (full regressions). MAP-RIGHT's relationship to Intermediate was not anticipated but does not seem unreasonable when it is recalled that spatial scanning is analogous to scanning a page of text for understanding. Intermediate involves tasks of reviewing or scanning accounting situations and accounting principles to provide guidance on appropriate accounting treatment of transactions. The relationship of MRT-% to Cost in partial regressions but not full regressions indicates that the grade variance explained by VSA in partial models is explainable by a covariate in the full models for that course.

In examining course differences, covariates become increasingly nonsignificant over the three samples, supporting the notion of underlying sample differences. In the primary sample, all covariates are related to Principles I performance. In the secondary sample, all covariates are related to Principles I performance, but only SAT-MATH, SAT-VERBAL, and GPA are related to Principles II performance. For the tertiary sample, no covariates except GPA are significant for any of the four courses.[33]

Although gender differences were only hypothesized for Principles I, they are also found in the course differences. Table 45 summarizes gender differences across courses. For the subjects continuing to Principles II, partial regression results indicate that VSA is related only to Principles I for females and only to Principles II for males. In full regressions, VSA is related to performance only for females and only for Principles I.

For the tertiary sample of subjects continuing to both Intermediate and Cost, course differences in the VSA-performance relationship are found only for females. MRT-% is related to females' Principles I performance. No VSA variable is related to performance in Principles II, Intermediate, or

---

[33]It should be recalled that SAT-MATH, SAT-VERBAL, and GPA are static measures and were all measured before the start of Principles I. GPA, for instance, reflects student performance in the semester prior to Principles I and is not updated for subsequent performance.

Cost. MRT-RIGHT's significance for Principles I (HW/QUIZ) in the secondary sample and nonsignificance for Principles I in the tertiary sample is consistent with the restriction of tertiary sample analysis to EXAMS only.

Table 45
Summary of separate analyses by gender
of course differences in the VSA-performance relationship

Secondary sample (N = 342):

	Significant relationships: Partial regression, females	Partial regression, males	Full regression, females	Full regression, males
Prin I	MRT-RIGHT (hw)	--	MRT-RIGHT (hw)	--
	MRT-% (exams,hw)	--	MRT-% (hw)	--
Prin II	--	MAP-RIGHT (exams)	--	--
	--	MRT-% (exams)	--	--

Tertiary sample (N = 71) (only EXAMS analyzed):

	Significant relationships: Partial regression, females	Partial regression, males	Full regression, females	Full regression, males
Prin I	MRT-%	--	--	--
Prin II	--	--	--	--
Inter.	--	--	--	--
Cost	--	--	--	--

The performance scores related to VSA. The performance scores for which VSA-performance relationships are found are EXAMS and HW/QUIZ. HW/QUIZ is significantly related to VSA more frequently for females than for males, while EXAMS are more frequently related to VSA for males than for females. MAP measures, when significant, are related to EXAMS only. MRT measures are related to EXAMS and HW/QUIZ.

COMPUTER is rarely related to VSA measures. This is consistent with the Ruf, Brown and Crawford (1991a) contention that accounting software tasks are often designed in such a way as to obviate the need for VSA. In that study, the group exposed to accounting software also had no MAP relationships, while the lecture-only group did show relationships between MAP and performance. Accounting software such as BGL imposes structure on processing tasks which may make VSA irrelevant.

The presence of a computerized practice set in the current research may also be responsible for the few MAP relationships found. That is, the COMPUTER use throughout the semester may have provided ways of visualizing the accounting information system which carried over to student performance on EXAM and HW/QUIZ, obviating the need for the aspect of VSA measured by MAP.

Interpretation of VSA measures. Research which examines the relationship of VSA to academic performance has generally used the number right on VSA tests as an independent variable. The current study analyzed additional variables--the number wrong and the percentage right--since these variables should logically contain information which is not contained in the number right.

The number wrong, however, may be difficult to interpret. MAP-WRONG and MRT-WRONG, in addition to being VSA error measures, may reflect a student's willingness to risk answering uncertain items, interest in task completion, lack of understanding of item instructions, lack of attention to task requirements, or emphasis on speed over accuracy. Some of these explanations would be expected to be positively correlated with performance, while others would be expected to be negatively related. If both positive and negative influences on performance were being measured by these variables, few or inconsistent relationships might be found. The current study does not provide data which would address these issues. However, the frequent significance found for MRT-WRONG suggests that it effectively captures an inability to perform certain VSA tasks.

The interpretation of MAP-% and MRT-% appears more straightforward. Whereas MAP-RIGHT and MRT-RIGHT probably reflect VSA processing speed, the percentage measures logically reflect accuracy on the MAP and MRT.

Results indicated fewer MAP relationships relative to the MRT relationships. This suggests that the visual scanning subfactor of VSA is not as important as the spatial relations factor. The repeated significance of MRT measures suggests that visualization may facilitate performance on performance in the first accounting courses.

Implications for future research. Further research could compare the effects on results of characteristics of different samples, sample sizes, course content, and test content and format. In addition to replicating the current research at other institutions, other research designs might prove fruitful.

An experimental research design might carefully control in-class behaviors of instructors and students as well as the content and format of exams, homework, quizzes and practice sets across sections and courses. To ensure a larger sample size for analysis of Principles II, Intermediate and Cost, the secondary and tertiary samples could be expanded by administering VSA testing to all students enrolled in each class instead of including only those originally tested in Principles I.

Qualitative analysis techniques could be applied to course content, in-class behaviors, and the requirements of exams, homework, quizzes, and practice set. For some sections, the in-class behaviors and grade requirements could be systematically varied. Lectures, homework problems and exams could be designed to maximize the need for VSA. Computer software with graphic capabilities could be explored for supplemental use in the classroom. Also, some sections of Principles I might include a module which encourages the use of visual-spatial techniques in thinking about accounting and solving accounting questions. Post-testing of VSA could assess the effect on VSA of such training as well as measure the effect on performance in the course. And, of course, the usefulness of different VSA instruments could be explored.

Research could also investigate other populations and behaviors, such as students in advanced undergraduate and graduate classes and accounting graduates taking the CPA exam. The VSA-performance relationship could be explored in the work context for various kinds of accountants. For instance, if VSA is helpful in problem-solving and constructing mental images of relationships, then it might very well be associated with competence at auditing tasks.

Contributions. The current research explored relationships not yet discussed in published literature--the role of VSA with completing, continuing and performing in the first four accounting courses offered in most curricula. This study investigated the VSA of students actually enrolled in accounting coursework to identify the VSA relationships which currently exist, that is, without experimental manipulation of course requirements. The analyses utilized VSA-only models as well as models which included covariates suggested in the literature. Relationships were identified for variables not previously explored in published research on VSA and academic performance--the number wrong and the percentage right on VSA tests.

The analyses of Principles I used a large sample to ensure a representative sampling of those students enrolled in the course at this university. The study was also a longitudinal one, tracking the original subjects through Principles II and on to Intermediate or Cost or both. This made possible, for a small sample, comparison of the role VSA played for these students in each of the four courses they completed.

In conclusion, this research suggests that VSA has some explanatory power beyond the aptitude measures traditionally used in admission and counseling decisions, as well as in explaining student behaviors related to accounting education. Unexplored aptitudes may be affecting student choices of academic major and coursework as well as behaviors once students are enrolled in accounting courses. With the recent emphasis on examining and improving accounting education and on encouraging problem-solving skills in accounting graduates, it is particularly important for accounting researchers to identify aptitudes and explore their relationships to accounting education issues.

REFERENCES

Ault, D.E., and M.R. Carver. 1987. The Effect of pedagogical format on the performance of MBA students in graduate-level accounting courses. Issues in Accounting Education: 161-172.

Baker, D.R. 1985. Predictive value of attitude, cognitive ability, and personality to science achievement in the middle school. Journal of Research in Science Teaching, 22, 2: 103-113.

Baker, J.C. 1990. Spatial strategies and spatial training in the collegiate curriculum. Paper presented at the Annual Meeting of the New England Educational Research Organization (Rockport, ME, May 2-4, 1990).

Baker, P.R., and J.C. Belland. 1988. Visual-spatial learning: Issues in equity. Symposium: Technological Equity: Issues in Ethics and Theory. Annual Meeting of the Association for Educational Communications and Technology (New Orleans, LA, January 14-19, 1988). ERIC Document # 295 626.

Baenninger, M. and N. Newcombe. 1989. The Role of experience in spatial test performance: A Meta-analysis. Sex Roles, 20, 5/6: 327-344.

Baldwin, B.A. and K.R. Howe. 1982. Secondary-level study of accounting and subsequent performance in the first college course. The Accounting Review, 57, 3: 619-626.

Battista, M.T. 1990. Spatial visualization and gender differences in high school geometry. Journal for Research in Mathematics Education, 21, 1, 47-60.

Benbow, C.P., and R.M. Benbow. 1984. Biological correlates of high mathematical reasoning ability. In Progress in Brain Research (Eds: G.J. De Vries et al.), 61 469-490.

Bergin, J.L. 1983. The Effect of previous accounting study on student performance in the first college-level accounting course. Issues in Accounting Education: 19-28.

Blough, P.M., and L.K. Slavin. 1987. Reaction time assessments of gender differences in visual-spatial performance. Perception and Psychophysics, 41, 3: 276-281.

Brown, M.L. 1990. The Relation of visuospatial and analytical skills and span of short-term memory to academic achievement in high school geometry. (Master's thesis, Virginia Polytechnic Institute and State University.)

Buckless, F.A., M.G. Lipe and S.P. Ravenscroft. 1991. Do gender effects on accounting performance persist after controlling for general academic aptitude? Issues in Accounting Education (Fall), 6, 2: 248-261.

Canlar, M. 1986. College-level exposure to accounting study and its effect on student performance in the first MBA-level financial accounting course. Issues in Accounting Education (Spring): 13-23.

Carpenter, P.A., and M.A. Just. 1986?. Spatial ability: an information processing approach to psychometrics. In Advances in the Psychology of Human Intelligence (Ed: R.J. Sternberg, Yale University). Hillsdale, New Jersey: Lawrence Erlbaum Assoc., Publisher: 221-253.

Carroll, J.B. 1985. Domains of Cognitive Ability. Paper presented at the meeting of the American Association for the Advancement of Science, Los Angeles. As reproduced in Snow, R.E., 1986, Individual Differences and the Design of Educational Programs, American Psychologist, 41, 10: 1029-1039.

Carter, C.S., M.A. LaRussa, and G.M. Bodner. 1987. A Study of two measures of spatial ability as predictors of success in different levels of general chemistry. Journal of Research in Science Teaching, 24, 7: 645-657.

Doran, B.M., M.L. Bouillon, and C.G.Smith. 1991. Determinants of student performance in accounting Principles I and II. Issues in Accounting Education (Spring), 6, 1: 74-84.

Dunn, W.M., and T.H. Hall. 1984. An empirical analysis of the relationships between CPA examination candidate attributes and candidate performance. The Accounting Review (October), 59, 4: 674-689.

Eckel, N., and W.A. Johnson. 1983. A Model for screening and classifying potential accounting majors. Journal of Accounting Education, 1, 2: 57-65.

Edmonds, T.P., and R.M. Alford. 1989. Environmental complexity and the level of information processing by introductory accounting students. Issues in Accounting Education (Fall), 4, 2: 345-358.

Egan, D.E., and L.M. Gomez. 1985. Assaying, isolating and accommodating individual differences in learning a complex skill. In Individual Differences in Cognition. R. Dillon (Ed). NY: Academic Press.

Ekstrom, R.B., J.W. French, and H.H. Harman. 1976. Manual for Kit of factor-referenced cognitive tests. Princeton, NJ: Educational Testing Service.

Elikai, F., and J. Baker. 1987. Empirical evidence on the effectiveness of quizzes as a motivational technique. Issues in Accounting Education: 248-254.

Eliot, J., and I.M. Smith. 1983. An International Directory of Spatial Tests. Windsor, Berks: The Nfer-Nelson Publishing Co., Ltd.

Eskew, R.K., and R.H. Faley. 1988. Some determinants of student performance in the first college-level financial accounting course. The Accounting Review, 63, 1: 137-147.

Fennema, E.H., and J.A. Sherman. 1978. Sex-related differences in mathematics achievement and related factors: a further study. Journal for Research in Mathematics Education (May): 189-203.

Friedlob, G.T., and R.M. Cosenza. 1979. Exploratory analysis for developing a predictive model of competence in accounting principles. In American Institute for Decision Sciences, Proceedings, 11th, New Orleans.

Gomez, L.M., D.E. Egan and C. Bowers. 1986. Learning to use a text editor: Some learner characteristics that predict Success. Human-Computer Interaction, 2: 1-23.

Hair, J.F, Jr., R.E. Anderson, and R. L. Thatham. 1987. Multivariate Data Analysis with readings. NY: MacMillan Publishing Co.

Hicks, D.W., and F.M. Richardson. 1984. Predicting early success in intermediate accounting: the Influence of entry examination and GPA. Issues in Accounting Education: 61-67.

Hyde, J.S. 1981. How large are cognitive gender differences? A Meta-analysis using w^2 and d. American Psychologist, 36, 8: 892-901.

Kail, R., and J.W. Pelligrino. 1985. Human Intelligence: Perspectives and Prospects. New York: W.H. Freeman and Co.

Kyllonen, P.C., D.F. Lohman and R.E. Snow. 1984. Effects of aptitudes, strategy training, and task facets on spatial task performance. Journal of Educational Psychology, 76: 130-145.

Linn, M.C., and A.C. Petersen. 1985. Emergence and characterization of sex differences in spatial ability: A Meta-analysis. Child Development, 56: 1479-1498.

Linn, M.C., and S. Pulos. 1983. Aptitude and experience influences on proportional reasoning during adolescence: focus on male-female differences. Journal for Research in Mathematics Education, 14, 1: 30-46.

Lipe, M.S. 1989. Further evidence on the performance of female versus male accounting students. Issues in Accounting Education (Spring), 4, 1: 144-152.

Lohman, D.F. 1989. Human intelligence: an Introduction to advances in theory and research. Review of Educational Research (Winter) 59, 4: 333-373.

Lohman, D.F, and P.C. Kyllonen. 1983. Individual differences in solution strategy on spatial tasks. In Individual Differences in Cognition, 1. R.F. Dillon and R.R. Schmeck (Eds). NY: Academic Press: 105-135.

Lord, T.R. 1987. A Look at spatial ability in under-graduate women science majors. Journal of Science Teaching (November): 757-767.

Lunneborg, C.E., and P.W. Lunneborg. 1984. Contribution of sex-differentiated experiences to spatial and mechanical reasoning abilities. Perceptual and Motor Skills, 59: 107-113.

Maccoby, E.E., and C.N. Jacklin. 1974. The Psychology of sex differences. Stanford, CA: Stanford Univ. Press.

McGee, M. 1979. Human spatial abilities: Psychometrics and environmental, genetic, hormonal and neurological influences. Psychological Bulletin (May): 889-919.

McNeill, E., and F. Collins. 1975. Personality tendencies and learning modes in elementary accounting. The Accounting Review (October): 888-897.

Meece, J.L., J.E. Parsons, C.M. Kaczala, S.B. Goff, and R. Futterman. 1982. Sex differences in math achievement: Toward a model of academic choice. Psychological Bulletin, 91, 2: 324-348.

Mumaw, R.J., and J.W. Pelligrino. 1984. Individual differences in complex spatial processing. Journal of Educational Psychology, 76, 5: 920-939.

Mutchler, J.F., J.H. Turner, and D.D. Williams. 1987. The Performance of female versus male accounting students. Issues in Accounting Education (Spring): 103-111.

Newcombe, N., M.M. Bandura and D.G. Taylor. 1983. Sex differences in spatial ability and spatial activities. Sex Roles, 9, 3: 377-386.

Ott, R.L., D.S. Deines, and D.P. Donnelly. 1987. The Use of a fundamental practice set in intermediate accounting. Issues in Accounting Education: 131-138.

Pallrand, G.J., and F. Seeber. 1984. Spatial ability and achievement in introductory physics. Journal of Research in Science Teaching, 21, 5: 507-516.

Pepin, M., R. Beaulieu, R. Matte, and Y. Leroux. 1985. Microcomputer games and sex-related differences: Spatial, verbal and mathematical Abilities. Psychological Reports (June): 783-786.

Poole, C., and G. Stanley. 1972. A Factorial and predictive study of spatial abilities. Australian Journal of Psychology, 24, 3: 317-320.

Pribyl, J.R., and G.M. Bodner. 1987. Spatial ability and its role in organic chemistry: a Study of four organic courses. Journal of Research in Science Teaching, 24,3:229-240.

Rebele, J.E., D.E. Stout, and J.M. Hassell. 1991. Review of empirical research in accounting education. Journal of Accounting Education (Fall), 9: 167-231.

Ruf, B., R.M. Brown, and H. Crawford. 1991a. Computer-assisted versus traditional accounting instruction: Differential moderating influences of math, verbal and visual-spatial abilities on performance. Working paper. Virginia Tech.

Ruf, B, R.M. Brown, and H. Crawford. 1991b. Cognitive determinants of learning accounting concepts: an evaluation of gender differences. Working paper. Virginia Tech.

Schroeder, N.W. 1985. The Effects of pre-college accounting on the college accounting student. Journal of Business Education (February), 60, 5: 207-211.

Schroeder, N.W. 1986. Previous accounting education and college-level accounting exam performance. Issues in Accounting Education (Spring): 37-47.

Snow, R.E. 1980 Aptitude processes. In Aptitude, Learning and Instruction (Eds: Snow, Federico and Montague). Hillsdale, N.J.: L. Erlbaum Associates.

Snow, R.E. 1986. Individual differences and the design of educational programs. American Psychologist (October), 41, 10: 1029-1039.

Snow, R.E., and D.F. Lohman. 1984. Toward a theory of cognitive aptitude for learning from instruction. Journal of Educational Psychology, 76, 3: 347-376.

Tyson, T. 1989. Grade performance in introductory accounting courses: Why female students outperform males. Issues in Accounting Education (Spring), 4, 1: 153-160.

Vandenberg, S.G., and A.R. Kuse. 1978. Mental rotations, a group test of three-dimensional spatial visualization. Perceptual and Motor Skills (October): 599-604.

Vicente, K.J. 1987. Accommodating individual differences in searching a hierarchical file system. (Master's thesis, Virginia Polytechnic Institute and State University.)

Wolff, A.S, and P.W. Frey. 1984-5. A Note on the correlation of visual-spatial ability and the acquisition of game skill. Educational Research Quarterly, 9, 4: 4-5.

Appendix A

ON THE OPSCAN FORM, PLEASE FILL IN (A) YOUR SUBJECT # STARTING WITH THE LEFT-MOST BLANK IN THE OPSCAN ID NUMBER BOX AND B) THE NUMBERS ON THE APPROPRIATE LINES FOR YOUR ANSWERS TO THE QUESTIONS. FOR EACH QUESTION, PLEASE DARKEN ONLY ONE ANSWER ON THE OPSCAN.

OPSCAN QUESTION

line 1 Your academic rank as of September 1991:
1 = freshman 2 = sophomore 3 = junior
4 = senior 5 = grad student
6 = other: ______________________

Use the choices below to answer the next two questions concerning your major and minor areas of concentration.

1 = Accounting 2 = Agricultural Economics
3 = Business 4 = Economics
5 = Finance 6 = Management
7 = Management Science 8 = Marketing
9 = none 10 = Other: __________

line 2 Your major area of concentration
line 3 Your minor area of concentration

---

line 4 Gender: 1 = Female 2 = Male

Age (use lines 4 & 5 to record the two digits in your age)
line 5 = first number in your age
line 6 = second number in your age (choice 10 signifies 0)

---

line 7 Is English your primary language or second language?
1 = Primary 2 = Secondary

---

line 8 Have you been enrolled in Tech's Acct 2004 before?
1 = No
2 = Yes, but I dropped before the drop date.
3 = Yes. I made a C so I am taking the course again.
4 = Yes. I made a D so I am taking the course again.
5 = Yes. I made an F so I am taking the course again.
6 = Other: ______________________________

---

line 9 Have you taken other accounting or bookkeeping courses?
1 = No.
2 = Yes, one course in high school
3 = Yes, more than one course in high school
4 = Yes, one course at another college or university
5 = Yes, more than one course at another university

OPSCAN QUESTION

line 10 Have you had any work experience in bookkeeping or accounting?
1 = No
2 = Yes, less than 6 months full-time.
3 = Yes, more than 6 months full-time.

---

line 11 What is the highest level mathematics you have passed? (in high school or univarsity)

1 = Algebra 2 = Trigonometry
3 = Calculus 4 = Calculus with Trig
5 = Calc with Matrices
6 = Other:______________________________

---

line 12 What is the PRIMARY reason for your taking Accounting 2115?

1 = This course is required for my major.
2 = I am taking this course as an elective.
3 = This course is a prerequisite for other courses I want.
4 = I am trying to get into the School of Business.
5 = Other:______________________________

---

line 13 What grade do you expect to be able to get in this course?

1 = A 2 = B 3 = C 4 = D
5 = pass on Pass/Fail 6 = Audit 7 = F

---

line 14 Which adjective best describes your feeling about taking this course?

1 = Excited 2 = Interested 3 = Mildly curious
4 = Neutral 5 = Resigned 6 = Worried
7 = Resentful

---

line 15 How many hours are you taking this semester?

1 = 3 2 = 4 - 6 3 = 7 - 9
4 = 10 - 12 5 = 13 - 15 6 = 16 - 18
7 = 19 - 21 8 = more than 21

---

line 16 Which describes your employment status this semester?

1 = I am not working this semester.
2 = I am working fewer than 10 hours per week.
3 = I am working 10 to 19 hours per week.
4 = I am working 20 to 29 hours per week.
5 = I am working 30 to 39 hours per week.
6 = I am working full-time (40 hours per week).
7 = I am working more than 40 hours per week.

WHEN YOU HAVE COMPLETED THIS PAGE, PLEASE STOP AND WAIT FOR INSTRUCTIONS. THANK YOU.

MAP PLANNING TEST — SS-3

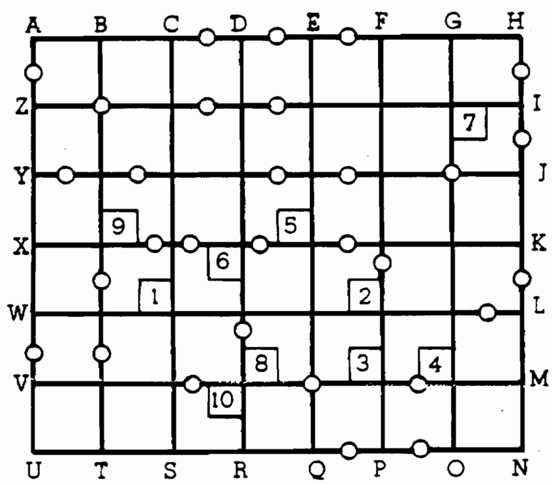
This is a test of your ability to find the shortest route between two places as quickly as possible. The drawing below is a map of a city. The dark lines are streets. The circles are road-blocks, and you cannot pass at the places where there are circles. The numbered squares are buildings. You are to find the shortest route between two lettered points. The number on the building passed is your answer.

<u>Rules</u>: 1. The shortest route will always pass along the side of <u>one and only one</u> of the numbered buildings.

2. A building is not considered as having been passed if a route passes only a corner and not a side.

3. The same numbered building may be used on more than one route.

Look at the sample map below. Practice by finding the shortest route between the various points listed at the right of the map. The first problem has been marked correctly.



The shortest route from:	Passes building:
1. A to Z	1
2. E to S	
3. P to J	
4. V to K	
5. O to F	
6. G to M	
7. D to Q	
8. F to T	

The answers to the other practice problems are as follows: 2 passes 5; 3 passes 3; 4 passes 2; 5 passes 4; 6 passes 4; 7 passes 6; 8 passes 5.

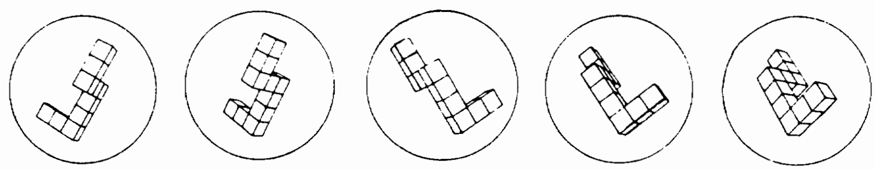
Your score on this test will be the number of right answers. It will <u>not</u> be to your advantage to guess unless you have some idea which route is correct. Work as rapidly as you can without sacrificing accuracy.

You will have <u>3 minutes</u> for each of the two parts of this test. Each part has one page. When you have finished Part 1, STOP. Please do not go on to Part 2 until you are asked to do so.

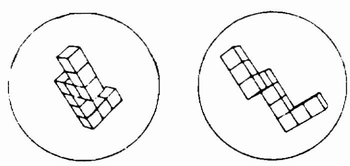
DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO.

M.R.T. TEST

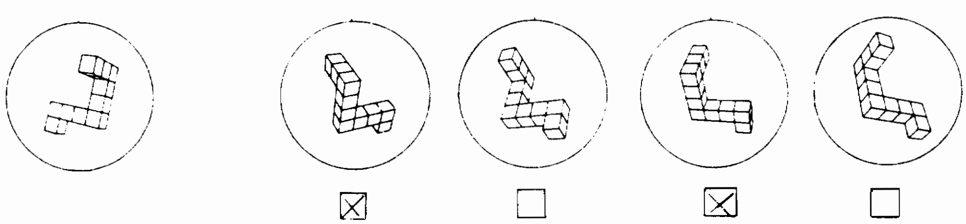
This is a test of your ability to look at a drawing of a given object and find the same object within a set of dissimilar objects. The only difference between the original object and the chosen object will be that they are presented at different angles. An illustration of this principle is given below, where the same single object is given in five different positions. Look at each of them to satisfy yourself that they are only presented at different angles from one another.



Below are two drawings of new objects. They cannot be made to match the above five drawings.
Satisfy yourself that they are different from the above.

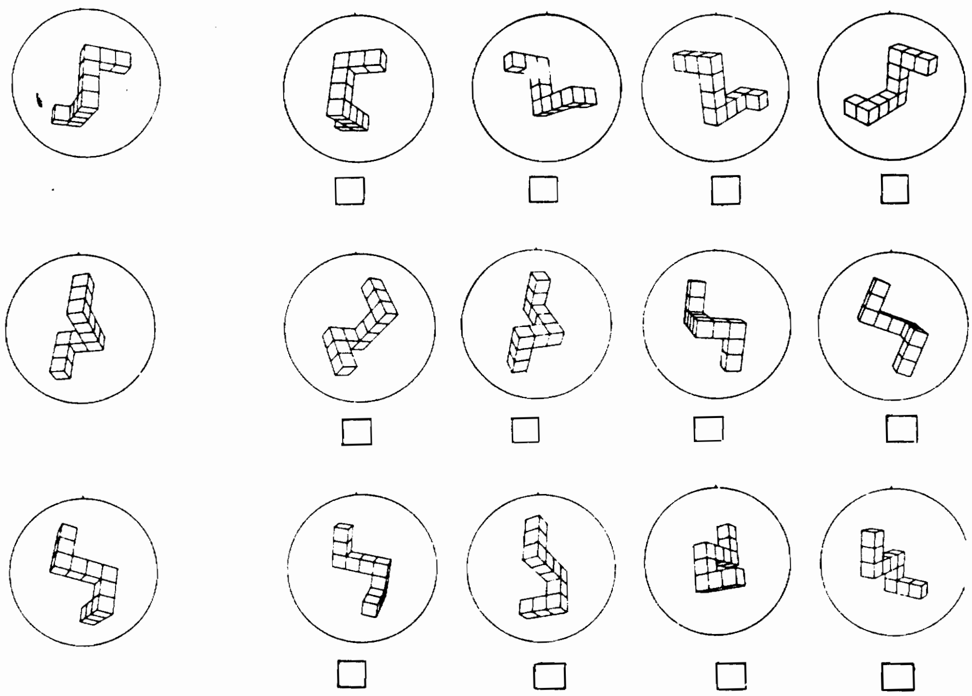


Now let's do some sample problems. For each problem there is a primary object on the far left. You are to determine which two of four objects to the right are the same object given on the far left. In each problem always <u>two</u> of the four drawings are the same object as the one on the left. You are to put Xs in the boxes below the correct ones, and leave the incorrect ones blank. The first sample problem is done for you.



Adapted by S.G. Vandenberg, University of Colorado, July 15, 1971
Revised instructions by H. Crawford, U. of Wyoming, September, 1979

Do the rest of the sample problems yourself. Which two drawings of the four on the right show the same object as the one on the left? There are always two and only two correct answers for each problem. Put an X under the two correct drawings.



Answers: (1) first and second drawings are correct
(2) first and third drawings are correct
(3) second and third drawings are correct

This test has two parts. You will have 3 minutes for each of the two parts. Each part has two pages. When you have finished Part I, STOP. Please do not go one to Part 2 until you are asked to do so. Remember: There are always two and only two correct answers for each item.

Work as quickly as you can without sacraficing accuracy. Your score on this test will reflect both the correct and incorrect responses. Therefore, it will not be to your advantage to guess unless you have some idea which choice is correct.

DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO

Vita

Dianna Ross Coker holds a Bachelor's degree in Philosophy from Hollins College in Virginia and a Master of Professional Accountancy from the University of Texas San Antonio. She is a Certified Public Accountant, licensed in the state of Texas.

Her business experience includes employment in corporate. public, and not-for-profit accounting. Ms. Coker was business manager and interim executive director of the YWCA in San Antonio, accountant for a Texas CPA firm, and manager of payroll operations for a national financial services firm.

Her teaching experience includes undergraduate and graduate accounting courses. Recent publications include legal research published in Oil and Gas Tax Quarterly and in Petroleum Accounting and Financial Management Journal. Ms. Coker is a member of the American Accounting Association and the American Institute of CPAs.

Dianna Ross Coker