

**EFFECTIVENESS OF A PROGRAM-SPECIFIC
ASSESSMENT INSTRUMENT FOR A
DEPARTMENT OF CLOTHING AND TEXTILES**

by

Valerie Littlejohn Giddings

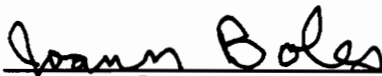
**Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements
for the degree of**

DOCTOR OF PHILOSOPHY

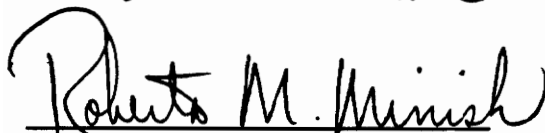
in

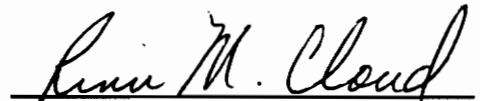
Clothing and Textiles

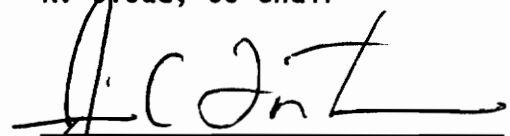
APPROVED:

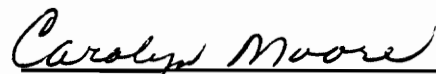

J. F. Boles, Co-chair


L. Andrew *RMC*


R. Minish


R. Cloud, Co-chair


J. C. Fortune


C. Moore

June, 1991
Blacksburg, VA

C.2

LD

5655

V856

1991

6533

C.2

**Effectiveness of a Program-Specific
Assessment Instrument for a
Department of Clothing and Textiles**

by

Valerie Littlejohn Giddings

(ABSTRACT)

The purposes of this study were (1) to determine the validity, reliability, and item quality of a program-specific, assessment instrument designed to measure student mastery of core knowledge in the discipline, (2) to evaluate the effectiveness of the instrument for determining student outcomes from the value-added perspective, and (3) to determine whether performance on the instrument and other measures of achievement was affected by student involvement characteristics.

The instrument contained 100 multiple-choice items relating to five core courses. For portions of the study, the instrument was divided into five subtests and administered as a pretest on the first day of classes in the five core courses and as a posttest on the last day of classes. The instrument in its full form was administered to upperclass clothing and textiles majors and to underclass clothing and textiles majors. Interviews related to the instrument were conducted with a subset of students who had taken the subtests and the full exam and with faculty who taught the courses.

The split-half and the KR-20 reliability estimates for internal consistency of the full test were above .70. The increase in student performance from pretest to posttest and the correlation between information obtained from the interviews with the results of the item

analyses demonstrated the instrument's content validity. The significant difference between scores received by upperclass majors and scores received by underclass majors on the instrument, and the correlations between course grades with scores on the instrument and final exam grades with scores on the instrument demonstrated the construct validity of the instrument. The ability of the instrument to measure value-added knowledge was ascertained by the increase in student performance from pretest to posttest.

Further data analyses revealed a pattern regarding factors which affect student achievement. Student involvement in organizations and employment had a positive and negative effect on various measures of student achievement.

Although the validity and reliability of the instrument were demonstrated, improvements in item quality and representation of course objectives on the core knowledge exam are needed.

DEDICATION

The author would like to dedicate this dissertation to her mother, Mrs. Ethel G. Littlejohn, whose love, support, and sacrifices over the years made this accomplishment possible. I would also like to dedicate it to my grandmother, the late Charity P. Littlejohn, who first introduced me to the field of clothing and textiles.

ACKNOWLEDGEMENTS

The author wishes to express her sincere appreciation to her advisors: Dr. Rinn Cloud and Dr. Joann Boles for their advice, encouragement, and constructive critique of the drafts. The research could not have been completed without the assistance of my committee members: Dr. Carolyn Moore for her advice and assistance with the data collection, Dr. Roberta Minish for her advice and encouragement throughout the research, Dr. Jimmie Fortune for his assistance with the overall methodology and statistical procedures, and Dr. Loyd Andrew for his advice and encouragement during the developmental stages which helped me to narrow the focus of the research and to complete the research in a timely manner. Although Dr. Andrew died shortly after my final exam he made valuable contributions through the entire process.

My gratitude is also extended to the faculty, staff, and students in the Department of Clothing and Textiles for their participation. I would also like to acknowledge the following individuals for their special contributions:

Dr. Marjorie Norton for her assistance and advice in locating a source of funding which allowed me to conduct my research full-time;

Ethel Bonds, my sister and editor, for generously giving her time and expertise to proof the many drafts;

Dove Robertson and Gina Eubanks for their love, encouragement, and expressions of confidence throughout this endeavor;

Janet Bryant for consistently reminding me of my goal and contributing her time and expertise with WordPerfect;
My sorors of Blacksburg Alumnae Chapter of Delta Sigma Theta Sorority, Incorporated for their support;
The members of First A.M.E. Church, Spartanburg, SC, for their love and support.

A very special thank you is extended to my husband, Ronald L. Giddings, and my son, Jarhon N. Giddings, for their love, patience (tolerance), and understanding from the first course taken to the final word written. Thanks also to my brothers and sisters - Ethel Bonds, Horace Littlejohn, William Littlejohn, Jacquelyn Davis, Wallace Littlejohn, and their spouses, who have loved, supported, and encouraged me throughout my educational experiences; their love makes this degree particularly meaningful. I would also like to recognize my father, the later Horace C. Littlejohn, Sr., whose memory continues to be inspirational in my educational and professional pursuits. Last, though definitely not least, I am sincerely grateful to my Father in Heaven and my Lord and Savior, Jesus Christ, for blessing this experience and the final product.

TABLE OF CONTENTS

ABSTRACT	ii
DEDICATION	iv
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vii
LIST OF TABLES	xv
CHAPTER I. INTRODUCTION & BACKGROUND	1
Introduction	1
Review of Literature	4
Background of the Current Assessment Movement	4
Mandates to Assess Student Performance	9
Demands from the Federal Government	9
Mandates from Accrediting Associations	10
Mandates from the State	10
Assessment of Undergraduate Education	12
Assessment to Improve Teaching and Learning	13
Assessment as a Threat	17
Assessment to Produce the Desired Results	18
Assessment Methods	20
Outcomes Measures	20
Advantages	20
Disadvantages	21
Models Using Outcomes Measures	23
Value-Added Measures	25

Table of Contents (Continued)

Instruments Used for Value-Added Measures	26
Advantages of Value-added Measures	27
Disadvantages of Value-Added Measures	28
Value-Added Models of Assessment	29
Measurement and Evaluation	32
Definitions	32
Research Studies Related to Criterion Validity	35
Test Development	41
Test Procedures	41
Types of Achievement Tests	42
Determining Test Quality	44
CHAPTER II. STATEMENT OF THE PROBLEM	46
Problem	46
Purpose	47
Theoretical Framework	47
Using Value-added Assessment	47
Characteristics Affecting Student Achievement	50
Demonstrating Validity and Reliability of Achievement	
Tests	51
Hypotheses	56
Pretest Influence	57
Criterion-related Validity	58
Construct Validity and Value-added Quality	58

Table of Contents (Continued)

Content Validity	58
Reliability	58
Item Quality	59
Student Involvement	59
Assumptions and Limitations	59
Limitations	59
Delimitation	60
Operational Definitions	60
CHAPTER III. METHODOLOGY	62
Description of the Department of Clothing and Textiles	62
Description of Core Courses	63
Development of Instrument	64
Description of the Final Instrument	65
Sampling Procedures and Data Collection	66
Analysis of Data	69
CHAPTER IV. RESULTS AND DISCUSSION	73
Description of the Sample	73
Subtest Sample	73
Consolidated Exam Sample	75
Subtest and Exam Results	84
Pretest Influence	86
Criterion-related Validity	88
Construct Validity	90

Table of Contents (Continued)

Content Validity	96
CT 1004 - Clothing and People	96
CT 2404 - Basic Apparel Assembly Lecture	98
CT 2414 - Basic Apparel Assembly Laboratory	99
CT 2604 - Introduction to the Fashion Industry	102
CT 3204 - Textile Science	105
Interview With Seniors	107
Summary	108
Reliability	108
Internal Consistency	110
Stability	112
Item Quality	112
CT 1004	114
CT 2404	117
CT 2414	119
CT 2604	121
CT 3204	123
Consolidated Core Knowledge Exam	125
Characteristics of Student Involvement	133
Chapter Summary	148
CHAPTER V. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	151
Summary	151
Pretest Influence	152

Table of Contents (Continued)

Criterion-related Validity	152
Construct Validity	153
Content Validity	154
Reliability	155
Item Quality	155
Characteristics of Student Involvement	156
Conclusions	157
Recommendations	160
Recommendations for Improving the Instrument	160
Recommendations for Using the Revised Instrument	161
Recommendations for Further Research	161
Recommendation from Advisory Committee	162
BIBLIOGRAPHY	164
APPENDIX A	174
Pretest Survey	175
Posttest Survey	176
Department of Clothing and Textiles Core Knowledge Exam	177
APPENDIX B	192
Protocol for Subtest Sample	193
Protocol for Instructors	194
Protocol for Seniors	195
APPENDIX C	196
Course Objectives and Related Subtest Items for Core Courses	197

Table of Contents (Continued)

CT 1004 - Clothing and People	197
CT 2404 - Basic Apparel Assembly Lecture	198
CT 2414 - Basic Apparel Assembly Laboratory	200
CT 2604 - Introduction to the Fashion Industry	202
CT 3204 - Textile Science	204
APPENDIX D	207
Item Analysis for CT 1004 Clothing and People	208
Item Analysis for CT 2404 Basic Apparel Assembly Lecture	210
Item Analysis for CT 2414 Basic Apparel Assembly Laboratory	212
Item Analysis for CT 2604 Introduction to the Fashion Industry	214
Item Analysis for CT 3204 Textile Science	216
APPENDIX E	218
Item Analysis for Core Knowledge Exam - Consolidated	219
APPENDIX F	227
Verification of the Key for the Clothing & Textiles Core Knowledge Exam	228
VITA	231

LIST OF TABLES

Table 1.	Summary of Data Analysis by Hypothesis	72
Table 2.	Frequencies of the Sample Participating in the Subtest Evaluations	74
Table 3.	Demographic Characteristics of the Total Subtest Samples	76
Table 3a.	Demographic Characteristics of the Pretest/Posttest Group	77
Table 3b.	Demographic of Pretest-Only and Posttest Only Groups . .	78
Table 4.	Additional Characteristics of the Total Posttest Sample	79
Table 4a.	Additional Characteristics of the Pretest/Posttest Group	80
Table 4b.	Additional Characteristics of the Posttest-Only Group .	81
Table 5.	Demographic Characteristics of Sample Taking the Consolidated Exam	82
Table 6.	Results of Subtests and Consolidated Exam	85
Table 7.	Comparison of Posttest Scores of the Pretest/Posttest Group and the Posttest-Only Group	87
Table 8.	Product-Moment Correlations Between Course Grades and Pretest Scores and Posttest Scores by Course and Over all Courses	89
Table 9.	Correlation Coefficients for Posttest Scores and Final Examination Grades	92
Table 10.	A Comparison of Pretest and Posttest Scores for the Pretest/Posttest Group	93
Table 11.	Comparison of Core Knowledge Exam Scores Based on Student Classification	95

List of Tables (Continued)

Table 12.	Course Objectives For CT 1004 - Clothing and People . .	97
Table 13.	Course Objectives for CT 2404 - Basic Apparel Assembly Lecture	100
Table 14.	Course Objectives for CT 2414 - Basic Apparel Assembly Laboratory	101
Table 15.	Course Objectives for CT 2604 - Introduction to the Fashion Industry	103
Table 16.	Course Objectives for CT 3204 - Textile Science	106
Table 17.	Summary of Instructor and Student Responses	109
Table 18.	Comparison of Reliability Estimates	111
Table 19.	Results of Test-Retest Stability Measure	113
Table 20.	Summary of Mean Difficulty and Discrimination Index For Subtests and Consolidated Exam.	115
Table 21.	Results of Item Analysis for CT 1004 - Clothing and People	116
Table 22.	Results of Item Analysis for CT 2404 - Basic Apparel Assembly Lecture	118
Table 23.	Results of Item Analysis for CT 2414 - Basic Apparel Assembly Laboratory	120
Table 24.	Results of Item Analysis for CT 2604 - Introduction to the Fashion Industry	122
Table 25.	Results of Item Analysis for CT 3204 - Textile Science	124

List of Tables (Continued)

Table 26.	Summary of Comparison of Item Analyses for Subtests and Consolidated Exam	127
Table 27.	Detailed Comparison of Item Analyses for Subtests and Consolidated Exam	128
Table 28.	Relationship Between Outcomes Scores and Number of Organizations for the Subtest Sample	135
Table 29.	Relationship Between Outcomes Scores and Numbers of Organization for Upperclass Majors	137
Table 30.	Relationship Between Outcomes Scores and Hours of Employment for Subtest Sample	139
Table 31.	Relationship Between Outcomes Scores and Hours of Employment for Upperclass Majors.	141
Table 32.	Relationship Between Outcomes Scores and Hours Spent on School-Related Activities for the Subtest Sample	144
Table 33.	Relationship Between Outcomes Scores and Hours Spent on School-Related Activities for Upperclass Majors	146

CHAPTER I

INTRODUCTION & BACKGROUND

Introduction

Private industry, state governments, the federal government, and the general public are seeking accountability in higher education. Well publicized reports assailing the quality of higher education have criticized the academy for not effectively preparing its graduates. Much of the criticism seems to center on the argument that too little attention is given to teaching and learning (Bok, 1986). Postsecondary institutions are therefore being pressured by state legislators and accrediting associations to assess and report the cognitive impact of the college experience on students.

In response many states require their public institutions of higher education to undertake assessment initiatives. The initiatives are also endorsed by accrediting associations who require assessment activities to be integrated into the self study of programs under review. As a result, the literature on assessment is replete with information regarding methods which have been used at various institutions. These methods vary, according to Ewell (1987a), in their focus on curriculum issues or on institutional issues.

Although institutions use a variety of methods to measure student outcomes, these methods can be categorized as outcomes measures and value-added measures. Outcomes measures demonstrate the knowledge,

skills, and attitudes of students upon graduation (exiting characteristics) but do not consider the entering characteristics. Value-added measures consider entering characteristics by pretesting students as freshmen and administering a post-test to determine the gains that can be attributed to the college experience. Value-added measures provide a better understanding of an institution's contribution to student development. The method selected, whether outcomes, value-added, or a combination of the two, depends upon the purpose of assessment (Terenzini, 1989).

For both methods, a variety of available instruments can be used. The instruments include commercially available standardized tests and surveys, professional licensing exams, and locally-designed exams and surveys. Although commercial instruments have been criticized for the general nature of their contents, they are convenient because their reliability has already been demonstrated (Ewell, 1987a). Program-specific instruments, however, may prove more valuable and the results more valid because they would be developed and evaluated by faculty within the field and would therefore be more closely tied to curriculum objectives. According to Ewell (1987a), these advantages and others make the use of program-specific tests more appealing than standardized tests. Most assessment authorities agree that successful assessment initiatives should be curriculum based and should use a variety of methods and instruments because of the diverse nature of learning in higher education.

Virginia Tech uses a variety of instruments to evaluate students at different levels of their college experience, including alumni status. The university tests students on a broad scale and requires departmental units to develop assessment procedures congruent with their missions and goals. The Department of Clothing and Textiles at Virginia Tech became involved in the university-wide testing of sophomores, seniors, and alumni in 1987. The department also began the development of an assessment instrument to measure student achievement of clothing and textiles core information. The instrument has been pretested on junior majors and once refined will be administered to seniors as a determinant of knowledge retention.

Purpose: The purpose of this study was to determine the validity and reliability of a program-specific, university-level, faculty-developed assessment instrument designed to assess student mastery of core knowledge in the discipline. Additionally, the purpose was to determine the effectiveness of the instrument for ascertaining student achievement from the value-added perspective and to determine whether performance was affected by student characteristics. The test, referred to as the core knowledge exam, was evaluated for internal consistency, stability, content validity, construct validity, criterion-related validity, item difficulty level, item discriminatory power, and item distractor effectiveness.

Review of Literature

Background of the Current Assessment Movement

The move for institutions to examine the effectiveness of undergraduate education is a result of several national reports criticizing the condition of the American educational system. The first report was the product of an in depth examination of education in the public school system conducted by the National Commission on Excellence in Education, a committee convened by then Secretary of Education Bell. "A Nation at Risk" (National Commission on Excellence in Education, 1983) described the condition of education in American public schools and provided recommendations for improvement. The findings of the commission included declining achievement of American students, a watered down high school curriculum and textbooks, and poor teacher preparation in terms of subject matter knowledge. These findings were supported by declining SAT scores, increasing numbers of functionally illiterate teenagers and adults, and complaints from employers that high school graduates lacked basic academic skills necessary to obtain entry level positions.

Since approximately fifty percent of high school graduates attend college, Bennett, former chairman of the National Endowment for the Humanities expressed concern about the quality of the institutions admitting these students ("To Reclaim a Legacy", 1984). Accordingly, the condition of higher education was the subject of other national reports. One such report was "Involvement in Learning" developed by the Study Group on the Conditions of Excellence in American Higher Education

("Text of New Report," 1984). The director of the National Institute of Education asked the group to assess the findings of "A Nation at Risk" as they related to postsecondary education. This assessment resulted in the report entitled "Involvement in Learning" which summarized the findings of their investigation and also included recommendations for improving the quality of undergraduate education.

Mortimer (1985), a member of the study group, reported evidence of the eroding quality of American undergraduate education. In his opinion, the evidence included declining "test scores of college graduates applying to graduate and professional schools" (p. 19), a decrease in the number of students declaring majors in the liberal arts and sciences, and a decrease in the rate of faculty salary increases. According to Mortimer, the report suggested that instead of using resources or inputs as a measure of quality, "all institutions of higher education should pay greater attention to the learning of the students they graduate as true measures of effectiveness" (p. 21). It also suggested the need for colleges to become more cognizant of research regarding the components of effective teaching and learning at the undergraduate level.

"Involvement in Learning" outlined three strategies which institutions could use to improve the quality of undergraduate education. First, the study group recommended that institutions increase student involvement in the learning process and increase contact between faculty and students. They also suggested establishing and achieving high expectations and standards (i.e., What are the

expected learning outcomes?). Finally the group recommended that institutions utilize assessment methods to make program improvements ("Text of New Report," 1984).

Another report, "To Reclaim a Legacy" (1984), which berated higher education, was developed by former Secretary of Education William Bennett. He criticized the general education curriculum of higher education for its lack of emphasis on the humanities. In his report Bennett described the findings of the National Institute for Education's Study Group on the State of Learning in the Humanities in Higher Education. The study group discovered that freshmen and sophomore level humanities courses were taught by graduate students and junior faculty, the humanities undergraduate curriculum was disintegrating, and the core curriculum of American colleges and universities lacked commonalities. As a result of these findings, the study group suggested that students would benefit more from experienced teachers teaching humanities courses. The study group also recommended humanities course requirements be spread out over four years rather than concentrated in the first two years of matriculation. According to the study group, the requirements should include a wide variety of in depth reading using primary sources of information.

Also mentioned in the report was the eroding quality of teaching in higher education, which was attributed to the emphasis on research in graduate school. According to Bennett ("To Reclaim a Legacy," 1984), "Many graduate schools have become so preoccupied with training narrow research specialists that they no longer address adequately the more

pressing need of higher education for good teachers, broadly versed in their fields, inspired by the power of their subjects and committed to making those subjects speak to the undergraduate" (p. 19). Bennett also urged university administrators, particularly the president, to become more involved in academic affairs and to place more emphasis on "good teaching in hiring, promotion, and tenure decisions" (p. 21).

Bennett's concern with college teaching and the curriculum was somewhat echoed in "Integrity in the College Curriculum," a report of the Association of American Colleges Committee (text reprinted in *The Chronicle of Higher Education*, February 13, 1985) which critically reviewed the undergraduate curriculum and the profession of college teaching. The committee characterized the curriculum as shallow, inconsistent, and lacking in structure. Students were found to be more interested in the credentials the degree afforded than the curriculum that constituted the degree. The report also criticized doctoral programs where the focus on research has led to the neglect of teaching. According to the committee, "only in higher education is it generally assumed that teachers need no preparation, no supervision, no introduction to teaching" (p. 29). Efforts to improve college teaching should be incorporated in doctoral programs, which should require courses regarding styles of learning and course methods and materials. Doctoral candidates, therefore, would be evaluated not only on the basis of their research skills but on their teaching skills as well. This emphasis on college teaching should also be apparent in the tenure and promotion process, and should further be recognized and promoted as a

viable profession by various associations such as the American Federation of Teachers and the Association of University Professors ("Integrity in the College Curriculum", 1985).

Others within the higher education community have recognized the lack in quality of teaching in higher education. Bok (1986) attributes this shortcoming to the fact that too little time and effort are aimed at improving learning. According to Bok, "The time faculties and administrators spend working together on education is devoted almost entirely to considering what their students should study rather than how they can learn more effectively or whether they are learning as much as they should" (p.18). Cross (1988) has noted that there is too little research related to whether teachers change in response to student performance on standardized tests. Warren (1987) suggests that faculty need to become more aware of what students are learning and use the information to adjust the instructional approach and/or the course requirements, thus equating faculty expectations with student potential.

In response to these concerns from the higher education community and the reports on higher education quality, external agencies urged institutions of higher education to develop assessment programs that would measure how much students are learning and to use the information to improve instructional programs. Moreover, state governing and coordinating boards, accrediting associations, and legislators have insisted that institutions of higher education adopt assessment.

Mandates to Assess Student Performance

Demands from the Federal Government

William Bennett, during his tenure as Secretary of Education, was very outspoken in his concern for the higher education curriculum. In a speech at Harvard University, Bennett emphasized two major points: the need for institutions of higher education to adopt a core curriculum which would focus on history and traditions of the Western world and the need for those institutions to use assessment to determine their impact on students ("Text of Secretary Bennett's Address," 1986). He also encouraged students to hold their universities accountable for their education. Bennett chastised American higher education for being more concerned with getting money from the federal government than focusing on curriculum reform and effectiveness. He suggested that institutions develop assessment criteria which would make the public aware of what the institution has contributed to the student's cognitive development. Implementation of this suggestion would require a change in the prevailing notion of quality since many colleges regard resources or input rather than gains in learning as a measure of quality (Mortimer, 1985).

Bennett's proposition highlighted one of the recommendations of "Involvement in Learning"; assessment should be "an organic part of learning" - "to provide information about the teaching and learning process and as feedback to help improve the effectiveness with which students, faculty, and the institutions carry out their work" ("Text of

New Report," 1984, p. 45). Accrediting associations have supported this notion of more emphasis on learning.

Mandates from Accrediting Associations

The federal government proposed that accrediting associations include in their standards an emphasis on student assessment (Manning, 1988). Although the federal government cannot regulate higher education or accrediting agencies, all six regional accrediting associations have adopted criteria related to student outcomes (Banta & Moffett, 1987).

The Southern Association of Colleges and Schools expects institutions under review in its region to determine program effectiveness using assessment and suggests procedures within its criteria which may be used. The association encourages institutions to use the results for program improvement (Southern Association of Colleges and Schools, 1989-1990). The North Central Association accrediting body has taken a different approach to assessment by encouraging institutions in its region to utilize existing standards and criteria to augment assessment programs (Ewell, 1987b). Hagerty and Stark (1989) discovered that professional accrediting associations in ten professional fields did not explicitly stress student outcomes in their standards. They concluded that "accreditors [should] stress curriculum dimensions rather than outcomes" (p. 18) since outcomes are usually based on a local rather than a national curriculum.

Mandates from the State

Legislators have also taken action on the assessment issue. The National Governors' Association published a report on education entitled

"Time for Results: The Governors' 1991 Report on Education" (National Governors' Association, 1986). The report encompassed several aspects of education, but the one most relevant is the section on college quality. In this section, the report discusses the findings of study groups regarding the effectiveness of higher education. The report strongly supports the state's role in requiring institutions to assess student learning and provides suggestions for assessment components. In addition, the report identified several state coordinating boards that have assessment requirements either already in place or in process. These states include Arkansas, Colorado, Florida, Illinois, Louisiana, Maryland, Missouri, New Jersey, New York, Ohio, South Dakota, Tennessee, Utah, and Virginia.

In Florida students are administered an examination during their sophomore year at the state's four year colleges and universities or during their last year at the community college (National Governors' Association, 1986). This examination is the equivalent of a rising junior examination. In New Jersey, the committee established by its state board recommended an assessment program with components which included assessing the general education curriculum and major program areas (Curry & Hager, 1987).

In 1985 the Virginia legislature requested information from the State Council of Higher Education pertaining to how student performance could be evaluated (Jaschik, September 1985). The state council responded with recommendations that all state-supported colleges and universities make plans to assess student achievement and report the

results to the council (Virginia Polytechnic Institute & State University Plan for Assessing Institutional Effectiveness, 1987).

Arkansas, Louisiana, and Illinois have made similar recommendations.

Other states, Colorado, Missouri, New York, New Jersey, Ohio, South Dakota, and Tennessee, relate assessment to funding. In Colorado, state universities must comply with assessment plans by 1990 or "lose 20 percent of their state funding appropriation" (National Governors' Association, 1986, p. 82). The mandate states "that institutions be held responsible for demonstrable improvements in student knowledge, capacities and skills between entrance and graduation" (Ewell, 1987b, p. 25). As an incentive for assessing student outcomes, the state of Tennessee provides a five percent supplement to the instructional budgets of universities that demonstrate program quality (Banta, 1985). Rather than prescribe assessment procedures, most states emphasize individuality in assessment plans based on the missions of the institutions (Ewell, 1987b). In a survey of 50 states, Boyer, Ewell, Finney, & Mingle (1987) discovered that two thirds of the state boards planned for some form of assessment. These boards felt an obligation to the public and legislators to ensure quality in their states' systems of higher education.

Assessment of Undergraduate Education

Mandates to adopt assessment encouraged by national reports have prompted institutions to develop and implement assessment initiatives. These initiatives, however, are reflective of the diverse nature of

higher education and vary according to the intended purpose of assessment.

Smith (1986) reports that the purpose may be to confirm student and/or institutional performance, to make comparisons with other units within or external to the university, or to make program improvements. Assessment, therefore, may involve institutional or program performance measures.

Institutional performance measures evaluate all students upon entry or exit from the university. These measures may evaluate the cognitive or affective domain. Students are also evaluated on a small scale for program assessment; the evaluation is conducted by a particular sub-unit at the university. Whether assessment is performed at the institutional or program level, a variety of methods and instruments may be used to gather student data: data which can be used for program improvement to enhance teaching and learning.

Although assessment can be used to enrich teaching and learning, many academicians view assessment as a threat to the traditions of the academy. In the final analysis, though, external constituencies can be satisfied and the academy can benefit if strategies are utilized which are congruent with the goals of the institution.

Assessment to Improve Teaching and Learning

Because the community of higher education is essentially decentralized, it yields the primary responsibility for curriculum decisions to the faculty of departmental units. This responsibility includes teaching subject matter information and evaluating the outcomes

of particular courses. This is a form of program based assessment which is not new to higher education. According to Ewell (1985), however, current assessment initiatives center on the institution rather than on program areas. Program-based assessment is forthcoming. To be effective, it must involve faculty and the results must be used for instructional decisionmaking rather than for accountability (Halpern, 1987). If used in this manner, assessment can generate positive outcomes for program areas.

Cross (1988) discusses the value of small scale assessment in the classroom as opposed to large scale institutional assessment. Since learning occurs in the classroom, classroom assessment should be used more frequently. Classroom assessment can provide feedback to teachers on how well what students are learning matches with the intended purpose of the course. However, assessment should not stop at the evaluation stage. Ewell (1987b) urges institutions to use assessment to improve student performance rather than collect data that document student performance. As a result, institutions would endorse teaching and learning.

Involvement in teaching and learning is a concept strongly promoted by Astin (1985a). He criticizes the valutive purpose of assessment and promotes a more active approach which he refers to as talent development. The talent development approach uses feedback to increase the involvement of students and faculty members and to develop their talents as fully as possible. He identifies three approaches to teaching that are currently used in academia: the content method, the

resources method, and the individualized method. The content method, which emphasizes subject matter conveyed by lecturing, is most popular but lacks the active participation of students. The resources method assumes that given the appropriate materials, such as a well equipped library and distinguished faculty, student learning will occur. Distinguished faculty, however, are often recruited with the understanding that they will teach few if any classes. The individualized method, which Astin endorses, is an instructional approach which tailors teaching to the student through independent learning experiences, self paced instruction, and other approaches that actively involve the faculty and student in learning. This method is the nucleus of the talent development approach.

Astin explains that the talent development approach also applies to the development of faculty in the area of teaching (Richmond, 1986). He chides universities for paying more attention to the development of research skills than to teaching skills. In accordance with this view, Hawley (1985) asserts that educational reform should "focus on enhancing the competence of teachers and fostering their effectiveness . . . because schools can only be as good as those who teach in them" (p. 187). Hence, attempts to improve teaching will undoubtedly improve the institution.

According to Warren (1987), assessment at the institutional level should be based on classroom activities and should reflect the involvement of both students and faculty. Inasmuch as the primary function of any university is learning, Krueger & Heisserer (1987)

suggest universities can use assessment to bridge the gap between the goal and realization of learning. Others who support assessment believe it can be instrumental in certifying accountability to university constituents while subsequently improving program quality (Banta & Moffett, 1987). Banta and Moffett further note that assessment can be used to "develop strategies for improving student retention, . . . suggest the content for faculty development programs, and stimulate curriculum reform . . . " (p. 37).

Institutions involved in assessment report that faculty have developed greater interest in students and the teaching learning process and that faculty interact more with their colleagues (Vandament, 1987). Halpern (1987) affirms that assessment provides a more effective method for establishing improvement priorities and evaluating the effects of actions to accomplish improvement. This type of action has been reported at the University of Tennessee where faculty were involved in developing comprehensive examinations which guided changes in curriculum and instruction (Banta & Moffett, 1987).

Assessment can help institutions or hurt institutions. Institutions can benefit by complying with state mandates and using the opportunity to focus on the teaching and learning process and to improve program effectiveness. However, if institutions do not take the initiative, a central authority outside the university may develop its own assessment agenda. This agenda may not be in compliance with the goals of the institution or its program areas.

Assessment as a Threat

Even though assessment has positive implications in terms of directing more attention toward student learning, not everyone affiliated with higher education supports assessment. Critics of assessment view it as a threat to institutional diversity, funding, and institutional and faculty autonomy. For instance, Jaschik (September 1985), reflecting the opinions of faculty and administrators, believes interference by external agencies may lead to the establishment of common goals of higher education. Diversity, according to Turnbull (1985), is what has made the U.S. system of higher education the best in the world. Feeling diversity is a quality that must be maintained even though it will hinder the establishment of assessment programs, Ewell (1987b) shares Turnbull's view.

Others fear assessment will absorb existing funds for institutional purposes and will eventually lead to budget cuts and program retrenchment (Boyer, et al., 1987). Currently some states, such as Tennessee, provide funding for assessment initiatives - funds which some believe could support other goals.

Another concern is that the results will be misinterpreted and used to compare institutions based on a single indicator, such as rising junior tests or standardized comprehensive exit examinations. Smith (1986) likens the use of a single index for comparison purposes to how SAT scores have been published and used as measures of system effectiveness.

Assessment also raises the issue of administrator and faculty relationship as it pertains to the classroom. Traditionally faculty have had full authority over curriculum matters. Because the current assessment movement may threaten that right, administrators are afraid of interfering with the professor's right to academic freedom and are also reluctant to tell faculty what or how to teach since it is believed that faculty are experts in their discipline area (Cross, 1986; Turnbull, 1985; Ewell, 1987b). These fears contribute to faculty and administrators' lack of enthusiastic involvement with assessment. This apprehension, however, is expected from institutions who have long held fast to notions of academic freedom, tenure, flexible work schedule, peer evaluation, collegiality, and other attributes of the academic community.

Assessment to Produce the Desired Results

Although assessment poses threats to the academy, if used effectively it can benefit students, faculty, the university, and external constituencies. According to Ewell (1987b) states are concerned only with the results of student performance and provide no recommendations for how assessment should be conducted. As a result, institutions have the freedom to design assessment strategies that are congruent with their missions and goals.

Assessment strategies which focus on program area effectiveness must include possible changes to what goes on in the classroom (Doyle & Hartle, 1985). According to Bok (1986), "A successful strategy of reform must enlist professors in individual institutions to work

together to improve not just the curriculum but the process of learning as well" (p. 21). Consequently, institutions need to base their assessment on some valid theory of learning (e.g. Astin's theory of student involvement). Astin (1985a) suggests that institutions need to focus less on what the institution does and more on what the students do. Quality, therefore, should be measured according to how effective an institution is in developing the talents of its students, the object of which is to enhance learning. Astin further claims that "the key to an effective learning experience is student involvement" (p. 36). His theory emphasizes that student learning increases with increasing quality and quantity of student involvement.

Since teaching links institutional goals with student learning, Boyer and others (1987) suggest that teaching must be given higher status if improvements in the undergraduate experience are to be realized. Boyer insists that the focus should be on better teaching, not on testing, if the quality of college education is to improve (Marchese, 1986).

State plans ignore the possibility that perhaps teaching methods should change and that this change may include adopting a particular theory which would ensure that learning takes place. Governing boards want data that illustrate how much students are learning at a given institution. This data can be provided in a number of ways without necessarily improving teaching and learning. Institutions, however, that want to take full advantage of the mandates and the funding

provided will attend to the examples set by others already involved in assessment.

Assessment Methods

Assessment methods can be grouped in two categories. The first category involves administering instruments on a one time basis to determine outcomes such as student knowledge or attitudes. The other category involves administering the same type of instruments more than once to determine the gains made from one administration to the other. These latter measures are referred to as value-added measures. The following section will discuss the two categories and provide examples of each.

Outcomes Measures

Traditionally, program quality in disciplines including law and medicine was based on the product of the educational process or the outcomes. Pace (1984a) notes that many people define outcomes as "whatever condition of knowledge or skill or development characterizes students at the conclusion of an education experience or program or at some later period in life as in alumni studies" (p. 14). These studies pertained to how well the graduates were able to get employment, what alumni salaries were, how well graduates were able to get into graduate school, or how students performed on exit exams.

Advantages. The advantages of outcomes measures are reflected in the perception of the public towards published information. Many schools use information relative to their graduates for public relations. For instance, the Virginia Tech Alumni Association 1986-1987

Annual Report reported the following: "Entering students are averaging over 1100 on the SAT's and have been high achievers in their respective high schools. They command good starting salaries when they have completed their work here. Many of our top students are going on to Graduate School" (p. 2). This is not a measure of the process itself - that is, what occurred during the four years - but merely an implication of what resulted from that experience. The essence of outcome measures is that good input (e.g. students with high SAT scores) contributes to good output and that the success of the graduate reflects the success of the program. This type of information is useful in attracting potential students to the university.

Another advantage of outcome measures is that they are not as expensive to administer as value-added measures. Outcomes measures alone are less expensive to implement and maintain since they require no additional testing or evaluation, other than perhaps followup of graduates.

Disadvantages. Outcomes measures ignore the probability that levels of student performance may differ when students enter school and would therefore differ when students complete four years of a college education (Astin, 1985b). They are measures of student quality rather than a measure of an institution's contribution to the student's level of performance. Astin (1985b) notes that the "outcomes level of competency is highly dependent upon its entering level of competency" (p. 95); thus, it is difficult to determine institutional impact. Yet, these measures are often used to judge the quality of institutions,

especially secondary schools. According to Tan (1986) "some researchers claim that highly productive students and successful alumni make a high quality department, but little empirical evidence is available to support this" (p. 25).

Outcomes measures, while serving a gate-keeping role, often prevent students from moving to higher educational levels. Astin (1987) therefore criticizes this method because it comes too late in the college experience to benefit students who may have needed extra assistance.

Outcomes measures are also closely linked to reputational measures of quality and, as Astin (1985b) states, "to an institution's position in the hierarchy. Thus, institutions at the top of the hierarchy - which tend to have good reputations, selective admissions, large endowments, and highly paid faculties - also manage to have low attrition rates and to produce large proportions of alumni who earn doctorates, make good salaries, and get listed in Who's Who" (p. 44).

Consequently, schools attempting to improve the performance of their graduates will take advantage of the outcomes measure by imposing more selective admissions policies (Spanghel, 1987), since, according to Astin (1985b), "the easiest way to enhance outputs is to enhance inputs" (p. 58). Selective admission policies, in Astin's (1987) opinion, will have a negative impact on minority access to higher education. Therefore, the increasing of admission standards by a large number of institutions would be detrimental to higher education since demographic and enrollment patterns indicate an increase in the number of minorities

who will make up the traditional pool of entering freshmen (Hodgkinson, 1983).

Models Using Outcomes Measures. Virginia Tech, in compliance with mandates placed by the State Council of Higher Education of Virginia, recently revised its mission statement and developed a plan for assessing the undergraduate education experience. This plan is based on the outcomes method of evaluating quality and identifies five goals that would provide the framework for assessing outcomes:

1. To enhance the intellectual development of students
2. To enhance the expression and communication abilities of students
3. To enhance the interpersonal, leadership, and organizational abilities of students
4. To enhance student knowledge and understanding of other cultures and international events
5. To enhance the occupational competencies of students. (VPI & SU Plan for . . ., 1987, p. 13)

The procedure for measuring the accomplishment of these goals includes administering, during the sophomore or junior year, an examination that would be based on the student's program of study, administering a senior exit survey (of student perceptions), and administering an alumni survey. The procedure also includes department-specific assessment methods. The university is intent on avoiding the use of standardized tests such as the ACT and the ACT-COMP because the results would provide little feedback to the departmental level, which is the primary focus of the plan. Therefore the university encourages the development of localized measures (VPI&SU Plan for . . ., 1987).

Florida uses a common examination testing program which requires all students to take an entry-level course placement exam and an exit

examination (College-Level Academic Skills Test). Although this form of assessment has not been a popular model for other states, one school in the Florida system has received notoriety for its successes. Miami-Dade Community College, regarded as the best two-year college in America, discovered their students were poorly prepared when they entered the community college and therefore placed greater emphasis on remedial training (Heller, 1988). Miami-Dade's method of assessment prevents students from taking courses for which they are not ready, and if students do not pass the College-Level Academic Skills Test (CLAST), they are prevented from continuing their education at a four-year institution. It was this denial of a student's access to a college degree that caused faculty to initially reject the program. However, the realization of the permanency of the situation generated faculty participation and eventual support.

The model used in the Florida system exemplifies an assessment program that does not allow the faculty to exercise much freedom in the curriculum. The examination is administered by the administration, and the administration decides whether the student completes a two-year degree or continues for a four-year degree. The college began to design their curriculum around preparing students for the examination and therefore placed greater emphasis on teaching to the test. As a result of the CLAST requirement, students at Miami-Dade have begun to take their education more seriously and opt for more challenging courses and teachers. Faculty, however, seem concerned that CLAST is the only measure of a student's future success in college and believe it should

be combined with some other performance measure, because the current system is particularly harmful to minorities (Ciereszko, 1987). Since minorities generally perform less well on standardized tests, this single performance index could greatly reduce the number of minorities attending four-year colleges.

Outcomes measures report end-result data and has its benefits and shortcomings, as previously reported. Value-added measures, however, provide data that illustrate gains made during the educational process.

Value-Added Measures

According to Ewell (1985), "Assessing the effectiveness of a curriculum or program means more than simply looking at what students know when they leave. It means taking into account their prior knowledge and characteristics" (p. 33). This concept of value-added has been heavily promoted as the most effective method for assessing the institution's impact upon student growth. Value-added is often used synonymously with talent development and, according to Halpern (1987), denotes "either the educational gains attributable to having spent time at a university or the amount and type of talent that an individual has developed as a result of university attendance" (p. 6). Astin (1985b) presents a strong case for talent development or the value-added concept for evaluating excellence in his book Achieving Educational Excellence.

The value-added assessment model, according to Astin (1985b), would identify excellent institutions as those "that have the greatest impact . . . on the students' knowledge and personal development . . ." (p. 61). Impact would be determined by measuring student performance at

entry and upon graduation. Gains made between entry and exit would be attributed to the university.

Assessing students upon entering the university is crucial to the value-added concept. Otherwise it would be impossible to determine what, if anything, the institution has added to the student's level of competence upon graduation (Keller, 1970; Astin, 1987). Since the essence of the value-added concept is " what the student has learned while enrolled . . . , programs are judged on how much they add to a student's knowledge and personal development" (Conrad & Wilson, 1985, p. vi).

Value-added methods can be used to measure growth in the cognitive and affective domains. A college student's development in the affective domain would entail understanding oneself and others, developing interpersonal relationships, establishing personal goals, and having a sense of morality (Pace, 1984a). Cognitive growth as defined by Astin (1985b) reflects the following skills: "verbal ability, quantitative ability, analytical reasoning ability, the ability to synthesize, organizational and human relations skills and understanding of cultural diversity" (p. 170).

Instruments Used for Value-Added Measures. Value-added methods do not rely on particular measurement devices. A variety may be used, but the ideal would be to use the identical or similar device at entry and exit points. Instruments that may be used to measure cognitive development include the American College Testing (ACT) Freshman Entrance Exam, the Scholastic Achievement Test (SAT), the College Level Academic

Skills Test (CLAST) the Graduate Record Examination (GRE), and the ACT College Outcomes Measures Project (COMP) (Ewell, 1985). Development in the personal and social domain may be evaluated with such instruments as the ACE/Cooperative Institutional Research Program administered out of UCLA, the ACT Evaluation Survey Service, and the NCHEMS/College Board Student Outcomes Information Service. Another instrument in this domain is Pace's College Student Experiences Questionnaire which measures student involvement (Pascarella, 1986).

Advantages of Value-added Measures. Many academicians have favored value-added measures because they provide an effective appraisal of the undergraduate experience. According to Jaschik (October, 1985), "Information obtained . . . allows curricula to be redefined and directed toward making up students' deficiencies. Students also receive a concrete measure of their education" (p. 1). Moreover, the value-added method can ensure the preservation of academic standards. Astin (1985b) suggests that it "will maximize the number of students who manage to attain minimum performance levels and the margin of safety by which students surpass those performance levels" (p. 69). Others (e.g., the president of the University of Southern Colorado) who support the value-added method claim that it promotes critical thinking of the educational process and is instrumental in demonstrating accountability to the public (Jaschik, October 1985). Evidence proves that the value-added method allows an institution to illustrate that substantial educational gain has occurred, thereby rendering more credit to the institution (Balderston, 1970). Benefits identified by McMillan (1988)

include placing students in courses which will improve deficiencies identified during freshman testing, tailoring assessment to the mission of the institution, examining the longevity of student gains, using the information for curriculum improvement, and fostering faculty and student involvement and interaction. Despite the many benefits of the value-added method, it is not devoid of critics.

Disadvantages of Value-Added Measures. Several concerns have been raised about the value-added concept. For instance some believe that the emphasis on improvement may overshadow the need to measure and establish minimum competency requirements for graduation. Others are apprehensive about the concept, indicating the difficulty in measuring improvement and linking growth to instruction (Belcher, 1987; Taylor, 1985). Students are exposed to a variety of experiences during their college attendance, the classroom is only one. Moreover, Belcher (1987) suggests that emphasis on the growth or improvement factor will lead to a reversal in selective admissions policies, since "high level entering skills provide little room for growth" (p. 33). Schools seeking to obtain a high rating of quality will take advantage of the value-added approach by accepting more freshmen who are poorly prepared and perhaps have more to learn, in order to yield higher improvement scores (Spangehl, 1987).

Faculty also fear that results will be used to evaluate their performance (Jacobi, Astin, & Ayala, 1987). Jaschik (September, 1985) warns that the value-added method may also encourage teaching to the test. However, Astin (1985b) argues that if students are expected to

know the information on a particular test, then teaching to the test will impart that knowledge. And imparting knowledge is the purpose of teaching.

Value-Added Models of Assessment. The previous discussion of value-added assessment and the following examples of value-added assessment in practice support the contention that for assessment to benefit the learner the program must first have a clearly identified purpose. Other factors that would contribute to a successful assessment program are the involvement of administrators and faculty, the use of a variety of measures which reduces error in measurement, the design of a curriculum-based instrument, and the use of results to guide curriculum improvement (Loacker, Cromwell, & O'Brien, 1986; Terenzini, 1989). The value-added method has been the primary form of assessment at Northeast Missouri State University (NMSU) for the past fifteen years. NMSU's strategy is one of the purest forms of assessment because it attempts to measure input from every possible angle (Taylor, 1985) and focuses on improvements in student performance. Jaschik (October 1985) provided a synopsis of the plan utilized by NMSU:

All entering students are required to take either the American College Testing Freshman College Entrance Exam or the ACT College Outcome Measures Project. The test results [ACT-COMP] at the end of the sophomore year . . . show how much students have learned in the general education curriculum and come at a point when they still have time to make up deficiencies. To graduate, seniors must take a national test in their major field. (p.22)

The university has documented improved productivity of students and faculty. Faculty involvement was preceded by an awareness of poor

student performance on national exams. Initially, the faculty responded by accusing the students of being incapable and the tests as being invalid. However, consistent low performance on national examinations prompted faculty to restructure programs and increase their level of expectation (Osigweh, 1986). These changes have resulted not only in improved student performance in the classroom and on national exams, but also in greater student involvement with academic activities, greater enthusiasm among faculty, and more interaction between faculty and administrators regarding teaching (Krueger and Heisserer, 1987). Osigweh (1986) attributes the program's success to (1) timing - the institution as a whole was ready for change; (2) open communication between administrators, faculty, staff, and students; and (3) participation by practically all members of the university.

Value-added testing, as part of a performance funding program, was implemented by the state coordinating board in Tennessee in 1979 (Jaschik, September 1985). Tennessee uses performance funding to motivate institutions in the state to improve their programs (Banta & Moffett, 1987).

Along with Northeast Missouri State University, the University of Tennessee at Knoxville (UTK) is often recognized for its successful use of the value-added method. The ACT-COMP is administered to all students at UTK during their freshman year and senior year. The scores, which measure growth, are interpreted by faculty and used to improve curriculum, teaching, and academic advising (Banta & Moffett, 1987; Jaschik, September 1985). The university's program succeeded in gaining

the involvement of faculty during the developmental stages. By 1984, the university reported change in teaching styles in response to a critical examination of the content of national exams and student performance on the exams. Many teachers switched from the traditional subject matter-lecture style to more emphasis on application of classroom learning with written assignments, assignments which enhance problem solving skills, term projects, field trips, and individual learning experiences (Banta & Moffett, 1987).

Teachers in the various disciplines developed comprehensive examinations which tested the competencies the faculty felt were representative of a graduate in their discipline. Faculty also openly discussed what was taught in their courses and how it was taught. Faculty then reached a consensus on their expectations of students and became more consistent in the teaching of core courses and course assignments. Value-added assessment at the University of Tennessee-Knoxville has also resulted in success in educational improvement across the board (Banta and Moffett, 1987).

A variety of testing instruments and procedures are used in the value-added method as described in the two situations in Missouri and Tennessee. The goals of their testing programs are to measure student gains and use the results to improve programs. The value-added method in these instances appears to have been successful in achieving the intended goals of assessment: improving undergraduate education.

Measurement and Evaluation

Whether an institution chooses the outcomes method, the value-added method, or both, some form of testing will occur. Therefore an understanding of measurement and evaluation theories as they pertain to test quality and test development is necessary. Payne (1974) asserts that measurement and evaluation are key elements in the instructional process. It is through measurement and evaluation that student achievement can be identified and adjustments in teaching can be made. Measurement and evaluation are valuable contributors to decision-making in education.

Definitions

Measurement, testing, and evaluation are terms often used interchangeably to imply similar meaning. However, experts in the field of measurement distinguish the three terms. Payne (1974) describes measurement as "the process of collecting, quantifying, and ordering information on an individual attribute, or object" (p. 3). Mehrens and Lehmann (1984) concur with this broad definition of measurement and describe evaluation as a subjective assessment of the results of the measurement process. According to Phillips (1968), steps in the evaluation process include the development of valid and reliable instruments to determine student achievement of educational objectives. Testing is the procedure of gathering the data. The results are then used to evaluate the learning situation.

A variety of methods can be used to obtain measurements in education. Testing is only one; others include rating scales and

observations. Tests are normally used to measure student cognitive development (Lindeman, 1967), which is best accomplished if there is an interaction between objectives, learning experiences, and measurement/evaluation activities (Payne, 1974). One form of testing is the use of achievement tests which Phillips (1968) suggests "provide a systematic way of observing students' responses related to the use of certain knowledge, skills, or understandings" (p. 38).

In order to be useful in educational decision-making, a measuring instrument must be valid, reliable, and useable (Lindeman, 1967). Validity refers to the ability of an instrument to measure what it has been designed to measure. Types of validity often mentioned in the literature are content validity, curricular validity, construct validity, concurrent validity, and predictive validity. Each type is appropriate for achievement testing. Content validity, however, is the most appropriate because it ensures that the instrument contains items which are representative of the instructional objectives, as well as the emphasis placed on the objectives during instruction. This type of validity uses the opinion of subject matter experts who evaluate the degree of correlation between the test and the instructional plan (Lindeman, 1967; McMillan, 1988). Curricular validity is a judgement regarding whether the test content was actually covered in class (McMillan, 1988; Mehrens & Lehmann, 1984). Construct validity as it relates to achievement testing is the extent the test distinguishes between levels of student achievement and measures a psychological trait. Construct validity of a test can be supported by its correlation

with other tests measuring the same construct, in addition to correlations with other student characteristics and environmental factors affecting performance (Lindeman, 1967).

Most researchers agree that concurrent validity and predictive validity are related to criterion-related validity because both types are correlated with a criterion measure. The difference between concurrent validity and predictive validity is based on the "time period when the criterion data are gathered" (Mehrens and Lehmann, 1984, p. 292). Predictive validity refers to the ability of a test to predict the future status of an individual while concurrent validity refers to whether the test score on a particular measure can represent a score on a similar measure (Lindeman, 1967). Construct validity and criterion-related validity of an achievement test are usually expressed with correlation coefficients.

Reliability is also a factor affecting the quality of a measuring instrument. Reliability relates to the consistency of the instrument to measure what is intended (Lindeman, 1967). Three procedures for determining reliability are 1) repeating the same test with the same group on two occasions; 2) administering equivalent test forms; and 3) dividing the test into two equivalent parts and computing a correlation coefficient.

Another factor affecting instrument quality is usability which considers such factors as time, cost, ease of administration, ease of scoring, and ease of interpretation (Lindeman, 1967). This factor often influences decisions regarding the test format, the development of

instructions, and whether or not to administer the test to a group or to individual students.

Throughout the literature in measurement there is consensus regarding the definition of validity and reliability, as well as agreement that both are important factors in test development. Although reliability is critical to validity, a test can be reliable without being valid. That is, a test can provide consistent results but may not be measuring the instructional objectives of concern.

Research Studies Related to Criterion Validity

The relationship between standardized test scores and school performance from kindergarten to professional school has been the topic of a number of research studies. The test scores are often used to predict future performance in school or on some other measure of achievement, which is a form of criterion-related validity. Most of the studies use statistical techniques such as simple correlations and multiple regression to test the hypotheses. The following discussion presents research relative to predictive validity, starting with subjects in elementary school.

Funk, Sturner, and Green (1986) investigated the ability of the McCarthy Scales of Childrens Abilities (MSCA) to predict achievement in kindergarten and in first and second grades. Achievement was measured by students' reading and math scores on subscales of the California Achievement Test (CAT). Pearson correlation coefficients revealed a strong relationship between the MSCA and the CAT reading and math scores. Multiple regression analyses indicated the general cognitive

index, a subscale of the MSCA, accounted for more variance than any combination of subscales. Another study which used the McCarthy scales also combined two other standardized tests, the Wechsler Intelligence Scale for Children (WISC-R) and the Woodcock-Johnson Tests of Cognitive Ability, to predict achievement from first to third grade (Reilly, Drudge, Owen, Rosen, Loew, and Fisher, 1985). Achievement was measured by the teacher ratings and scores on the Wide Range Achievement Test (WRAT). Using repeated measures ANOVA, the authors discovered the three tests were similar in their ability to measure student achievement. A correlational analysis revealed that each of the three tests was able to predict later achievement on the WRAT.

Massoth (1985) also used elementary school children as the subjects in a five-year followup study. He used the McCarthy Scales, administered during kindergarten, to predict sixth grade performance as measured by an average of student grades in five subject areas. Significant correlations were computed between three McCarthy Scales indexes (e.g. Quantitative, General Cognitive Index, and Memory) and average course grades.

Predictive validity studies involving high school students have also compared scores on achievement tests with actual performance in school. For example, Hanna, Sonnenschein, and Lenke (1983) administered the Orleans-Hanna Algebra Prognosis Test to eighth grade students in twenty states. The prognosis test and five other variables (grades in math, science, English, social studies, and a predicted grade in algebra) were used as predictors of final grades in algebra. Results of

a multiple regression procedure revealed the predictive validity of all six variables combined was greater than any of the variables alone, particularly the prognosis test.

Bloland and Michael (1984) also used the Orleans-Hanna Algebra Prognosis Test in a more complex study involving nine predictor (independent) variables and four criterion (dependent) variables. The study comprised ninth and tenth grade students enrolled in first year algebra to examine the relationship between final examination grades and course grades for two semesters (criterion variables), and scores on three achievement tests and chronological age (predictor variables). The achievement tests used in this study were six subscales of the Iowa Tests of Basic Skills, the Arlin Test of Formal Reasoning, the Orleans-Hanna Algebra Prognosis Test. The four criterion variables yielded significant correlations with most of the predictors. Results of zero-order correlations and multiple correlations revealed age and algebra prognosis were better predictors of final examination grades for both semesters; first semester grades were significantly correlated with algebra prognosis, the math concepts subscale, and the math computations subscale; and second semester grades were significantly correlated with algebra prognosis, mathematics concepts, and age.

Szymczuk and Frerichs (1985) investigated the ability of standardized tests to predict achievement in a high school computer course. Correlation coefficients and stepwise multiple regression analysis were used with four subtests of the ninth grade Iowa Test of Educational Development (ITED), five subtests of the Science Research

Associates' Computer Programmer Aptitude Battery, and the Wolfe Programming Aptitude Test-Form W that served as predictors. The dependent variables were students' first quarter, second quarter, and total point accumulation of classroom test scores. Two of the standardized test were administered during the first four days of class. The ITED scores were collected from student files. The authors demonstrated that the ITED Quantitative Thinking and Reading subtests were better predictors of student performance in the introductory computer course.

In research studies focusing on predictive validity, any number of independent or dependent variables may be used. To illustrate, Dong, Sung, and Dohn's (1985) study used seventeen independent variables (predictors) and ten dependent variables (criterion) to determine the relationship between subtests of the Ball Aptitude Battery (BAB) and grades in high school courses. This study also used correlations and stepwise multiple regression analysis, which indicated that the BAB vocabulary subtest was the best predictor across subjects.

The ability to predict success in college in specific courses or program areas is useful for recruiting, placement, and advising purposes. Hence, a number of studies have attempted to identify valid predictors of course performance. One such study was Gougeon's (1985) investigation of SAT mathematics scores as predictors of overall achievement in college math. Low to moderate correlation coefficients led the author to conclude that SAT mathematics scores do not predict achievement in college math. Another study which focused on predicting

course achievement used grades in each of two principles of accounting courses, and the student's score on the American Institute of Certified Public Accounts test (AICPA), a standardized accounting achievement test, as predictors of student success in Intermediate Accounting (Morgan, Cornick, and Kauder, 1985). The most valid predictor of success in Intermediate Accounting I was the average of the grades in Principles of Accounting I and II; the next best predictor was the AICPA test. Interestingly, the authors discovered that the best predictors for transfer students, when compared with non-transfer students, were the AICPA test score and the grade from Principles of Accounting II.

Hakstian and Woolsey (1985), interested in predicting overall achievement during the freshman year, used six tests of the Comprehensive Ability Battery (CAB), administered during the early part of the first semester. The criterion measures were final course grades in psychology, English, mathematics, biology, chemistry, physics, mean grade point average of the three science courses, and overall grade point average. The authors surmised, based on the results of correlational analyses, that the six CAB tests "possess validity for academic achievement criteria at the college and university level" (p. 340).

Similar studies have been conducted using subjects in graduate and professional school. The Graduate Record Examination is considered the "permit" to enter graduate school. Naturally, the predictive validity of GRE scores would be of interest to academicians in various disciplines. Consequently, Thornell and McCory (1985) conducted such a

study using three GRE scores (verbal, quantitative, and total) as predictors and graduate grade point average as the criterion variable. Correlations between total scores on the GRE and graduate grade point average were significant for the total sample and for each of the four disciplines included in the study. This result implies that the GRE is a valid predictor of success in graduate school.

Jones and Thomae-Forgues (1984) conducted an extensive investigation of the predictive validity of the six subtests of the Medical College Admission Test (MCAT). Predictor variables for the study were the six subtests, undergraduate grade point average, and a combination of the MCAT scores with undergraduate GPA; criterion variables were grades in the first and second years of medical school and total score on the National Board of Medical Examiners Part I exam. The results of correlation analyses revealed undergraduate GPA in science courses was the best predictor of medical school course grades and combined MCAT scores was the best predictor of performance on the MBME Part I exam. A similar study using the NBME Part III exam as the criterion measure was conducted by Calhoun, Wolf, Maxim, and Davis (1985). The predictor variables were undergraduate science and nonscience GPA, scores for the four components of a 1977 version of the MCAT, and mean interview faculty rating of medical school applicants. The results of the correlation analyses demonstrated a significant relationship between the faculty rating and two subtest scores of the MCAT - verbal ability and science.

Although the studies reported in this section have shown that in certain situations standardized tests can predict performance in elementary school, high school, college, and even graduate and professional school, the predictor variables do not account for all of the variance in the criterion measure; thus researchers need to consider as many independent variables as possible. Moreover, each of these studies did use multiple predictor variables which support Pascarella's (1986) contention that a number of factors in the college experience affect student performance. Although it is impossible to identify all variables that may affect student achievement, a thorough review of related literature is essential in identifying relevant variables.

Test Development

The major purpose of a test is to determine the effectiveness of teaching as reflected in student achievement of instructional objectives (Stanley and Hopkins, 1972). Reference to course objectives should be the primary concern of the evaluator during test development.

Therefore, instructional content should be arranged in outline form with areas and degrees of emphasis for particular objectives indicated. This outline is often referred to as a table of specifications which can serve as a guide during test development (Lindeman, 1967).

Test Procedures

Test procedures used most frequently in educational measurement from kindergarten through graduate school are oral exams, essay tests, and objective tests. Oral exams are traditional forms of assessment that have, since the turn of the century, been replaced by teacher-made

tests (Ebel, 1972). They are useful for indepth probing and evaluating the examinee's ability to express himself verbally in the presence of a group of people (Payne, 1974). However, scoring of oral examinations is highly subjective.

Essay tests allow the student more latitude in written expression and allow the evaluator to assess higher levels of cognitive abilities as well as basic composition skills, such as spelling, grammar, and punctuation. These tests, though valuable, are not as reliable as objective tests. The scoring is time consuming and subjective (Payne, 1974).

Objective tests utilize one or a combination of the following item formats: multiple-choice, short answer, true-false, or matching. The most common and flexible format is multiple-choice because it can measure various levels of cognitive outcomes such as knowledge, application, analysis, synthesis, and comprehension (Lindeman, 1967; Mehrens and Lehmann, 1984; Payne, 1974; Phillips, 1968). Scoring of objective tests is easier and more reliable than other procedures.

Types of Achievement Tests

Tests are also classified according to whether the data are used to compare student performance with a national sample or a local classroom sample. Achievement tests measure student learning in a given subject, allowing students to be ranked based on their level of knowledge. The two types of achievement tests are standardized tests and teacher-made tests. Standardized tests, as defined by Payne (1974) are instruments "that (1) contain a specified set of items, (2) include

explicit and standardized directions for administration and scoring, and (3) have usually been administered to a representative population of individuals for the purpose of securing normative data" (p. 90). These tests have been developed by measurement experts and usually have a high degree of reliability. There are three classifications of standardized tests - aptitude tests, achievement tests, and interest, personality, and attitude inventories (Mehrens and Lehmann, 1984). Although high in reliability, the content of standardized achievement tests is usually general in nature and therefore lacks relevance to local learning situations (Payne, 1974; Mehrens and Lehmann, 1984). Therefore, if the purpose for testing is to measure student achievement based on a particular curriculum, then teacher-made achievement tests would be more relevant and valid.

Teacher-made achievement tests and standardized achievement tests share two features: (1) they both measure student "knowledge and skill at a particular time" (Mehrens and Lehmann, 1984), and (2) they both utilize to a great extent an objective format for ease and objectivity in scoring (Payne, 1974). Teacher-made objective tests, which overcome the disadvantages of essay tests and oral exams, can be tailored to measure specific objectives in a particular curriculum. These tests can also be developed using a variety of formats - true/false, multiple-choice, matching, and completion (Ewell, 1987a; Phillips, 1968). A well constructed test can provide highly valid and consistent results that can subsequently be used to identify student strengths and weaknesses,

as well as areas in need of improvement in the curriculum (Ewell, 1987a).

Determining Test Quality

Most measurement and evaluation textbooks offer suggestions for writing effective items regardless of the format. The quality of the item determines the quality of the results which can be evaluated by item analysis procedures. Item analysis provides a detailed diagnosis of test results so items which are weak can be eliminated or improved and areas of student strengths and weaknesses can be identified (Payne, 1974).

Item analysis provides three classes of information: (1) the difficulty level of each item, (2) the ability of the item to distinguish between high achievers and low achievers, and (3) item response patterns to evaluate the distractors in multiple-choice items (Lindeman, 1967; Payne, 1974). Each procedure can be valuable for improving the reliability and validity of teacher-made achievement tests.

Ebel (1965) identifies ten factors which may be used to evaluate tests: relevance, balance, efficiency, objectivity, specificity, difficulty, discrimination, reliability, fairness, and speededness (reasonable test length for time allowed). Most of these are self explanatory and stress the necessity of basing test construction on course objectives and actual instruction, as well as ensuring that answers to items can be validated by experts in the field. The test should also be able to distinguish between levels of student knowledge.

Ebel and others (Payne, 1974; Stanley and Hopkins, 1972) agree that speed should not be a significant factor in achievement test scoring; a student's score should reflect his level of knowledge, not how fast he can complete the test.

In summary, measurement, evaluation, and testing are not interchangeable terms. The latter term provides the information necessary to conduct both measurement and evaluation processes. For these processes to be effective, the purpose of the test must be clearly defined in terms of the objectives to be measured, and these objectives should be closely related to the instructional process.

CHAPTER II

STATEMENT OF THE PROBLEM

Problem

Students attend college for a variety of reasons. But the goal for most students is to acquire in depth knowledge of a particular subject area. Whether this knowledge has been acquired is believed by some to be reflected in the student's overall quality credit average (QCA) or QCA in major. The recent emphasis on measuring the quality of the undergraduate experience has led many institutions to identify desired student outcomes and ways of measuring them. Since the majority of student time is spent in one program area, the development of assessment instruments related to instructional objectives and course content within the major is essential.

As part of Virginia Tech's total assessment plan, departments within the university were asked to submit individual plans for measuring student outcomes within their program areas. The Department of Clothing and Textiles submitted a five-year plan which included, among other initiatives, the development and use of an instrument that would measure student outcomes related to the department's core curriculum. The instrument has been developed and used on a trial basis, but the effectiveness of the instrument in measuring outcomes related to the curriculum and in measuring outcomes from a value-added perspective has not been established.

Purpose

The purpose of this study was to determine the validity, reliability, and item quality of a program-specific, university-level, faculty-developed assessment instrument designed to assess student mastery of core knowledge in the discipline. Additionally, the purpose was to determine the effectiveness of the instrument for ascertaining student outcomes from the value-added perspective and to determine whether performance was affected by student characteristics. The instrument was evaluated for internal consistency, stability, content validity, construct validity, and criterion-related validity, item difficulty level, item discriminatory power, and item distractor effectiveness.

Theoretical Framework

The framework for this study was based on the theory of talent development as it pertains to the use of value-added measures of student achievement and student characteristics affecting achievement. The study also encompasses the concepts of validity, reliability, and item quality. Discussions of the use of value-added assessment and student involvement, as well as the evaluation of assessment instruments using concepts of validity, reliability, and item quality will be presented in the following section.

Using Value-added Assessment

The value-added method for measuring student outcomes was discussed in depth in the previous chapter. The information presented

relative to the advantages and the models suggests this method is more effective than outcomes only for determining student achievement. Areas in which achievement can be measured are in the general curriculum and in a specific program area. Astin (1987), the major proponent for the value-added method offers six suggestions for conducting value-added assessment. First, the assessment results should be shared with the students to assist them in the classroom and in other areas of their experience. This sharing of information should help to increase faculty-student interaction. Second, institutions should use existing instruments, such as placement tests or classroom final examinations, as pretest and posttests to measure student knowledge change. Third, timing of test administration should be convenient for all involved. Fourth, objective measures using the multiple-choice format can be used to determine student growth by examining increases in the number or percentage of items answered correctly. Fifth, curriculum planning can be guided by evaluating student performance on individual test items to distinguish between the most and least challenging items. Sixth, Astin recommended that institutions begin new assessment programs "modestly and with minimal disruption of institutional activities" (p. 104).

Astin portrays value-added measurement as a relatively uncomplicated process involving a pretest and a posttest. Yet Pascarella (1986) suggests additional factors should be considered to determine what characteristics have influenced student change. According to Pascarella, many attempts at using value-added assessment have assumed that the impact of college is the same for all students -

regardless of different characteristics such as aptitudes, majors, and course taking patterns. Simple change on measures of student achievement can be misleading because not all students experience college in the same way. Therefore, Pascarella advocates causal modeling which allows one to investigate the various measures of the college experience with outcomes such as achievement tests, self-concept, or success in graduate school. This method permits a clearer understanding of how student development occurs (Pascarella, 1986).

An additional value-added technique involves testing freshmen and seniors using the same instrument and comparing scores of the two groups. However, Terenzini (1989) raises a concern regarding this type of cross-sectional research which he feels erroneously assume the two groups are similar. Therefore, Pascarella (1988) purports that the problem can be rectified by using SAT and ACT scores as control variables to match students. Similarly, McMillan (1988) feels the value-added method can be strengthened by making the following comparisons: comparing students attending different institutions, comparing resident and non-resident students, or comparing students in a given major with varying coursetaking patterns. Another alternative is to use longitudinal data, which, although prone to drop-outs in the sample, test wiseness, and statistical regression, can be enhanced by using a large sample at the beginning of the study (Pascarella, 1988). Still another useful method, according to Terenzini (1989), is the criterion-referenced approach to "examine the trend over time in the proportion of students who score above a faculty-determined threshold

of 'acceptable' performance on any given measure" (p. 661). Terenzini and McMillan both extol the benefits of gathering unit-specific data to guide program improvements.

The essence of value-added assessment is the change score representing student achievement based on a comparison of pretest and posttest results. Although supporters of assessment question the value of simple change scores, they have provided a number of useful suggestions for improving value-added research designs such as matching students on control variables or using longitudinal data.

Characteristics Affecting Student Achievement

Astin's (1985b) theory of student involvement indicates that the more students are involved in the college experience, the more benefits they enjoy in terms of satisfaction and achievement. Boyer (1987) discovered in a national survey of undergraduates that students were involved in a number of activities; this involvement varied in degree for different students. Activities which Boyer and Astin identified were paid employment (on and off campus), interaction with faculty, involvement in campus organizations, academic activities (studying), interactions with other students, and leisure activities such as reading and watching television. Both authors noted that an increasing number of students were involved in paid employment. Boyer suggests that although working can be beneficial to student achievement, it can also be detrimental if the number of hours is excessive. It is also interesting to note that even though Astin (1985b) implies excessive studying can negatively affect various aspects of the college

experience, Pace (1983) notes there is a positive relationship between student achievement and the amount of time spent on academic activities.

Demonstrating Validity and Reliability of Achievement Tests

The relationship between standardized test scores and school performance has been the topic of a number of studies previously mentioned in the review of literature. For the majority of the studies, the relationship of concern was the predictive ability of certain independent variables to determine the future status of an individual based on identified criterion variables. Since predictive validity is relevant to the current study, it is important to understand the variables and procedures used in similar studies.

The majority of the predictive validity studies used multiple regression and simple correlational analyses to determine the relationship between predictor variables and one or more criterion variables. In many instances the criterion variables were student performance in one particular subject area (Morgan, Cornick, & Kauder, 1985; Szymczuk & Frerichs, 1985), in several subjects (Hakstian & Woolsey, 1985; Dong, Sung, & Dohn, 1985; Hanna, Sonnenschein, & Lenke, 1983), or overall performance depicted by grade point average (Thornell & McCorry, 1985). The predictor variables included overall scores on achievement tests whose subject matter was similar to the criterion variable, scores on the separate subtests of the achievement test, or student predicted grades within a particular course. The authors reported administering the achievement tests (predictors) on the first

few days of class or using achievement test scores available from student files.

Construct validity, which refers to the ability of two instruments to measure the same construct, is also important to this study. Payne (1974) identifies five methods that may be used to demonstrate the construct validity of a test. These methods include (1) comparing groups who are known to differ on the variable of interest, (2) observing changes in student ability over a period of time, (3) comparing student performance on the test of interest with another test which measures the same construct, (4) evaluating the homogeneity of the test, and (5) observing the "test-taking process" (p. 258). Construct validity was demonstrated in Reilly et.al. (1985) when the authors successfully compared estimates of cognitive ability among three achievement tests. Repeated measures ANOVA was used to compare the mean scores on the three tests.

Few, if any, research studies have focused on content validity. This type of validity is a subjective assessment made by experts in a given discipline. In reference to a test, experts would be asked to match the content of the test to a set of objectives. A method used by Prevatt (1984) for demonstrating the content validity of an exemption test involved comparing pretest and posttest scores. This method also demonstrates construct validity.

The reliability of an instrument is normally reported to establish its credibility. According to Ebel (1972), reliability is affected by test length, item discriminatory power, item difficulty, item

homogeneity, and group variance. Reliability, therefore, is greater for long tests whose items have high discrimination indices (preferably above .30), whose items also have difficulty indices between .25 and .75, and with test scores from a group with high individual variance. Although many of the references consulted avoided identifying a cutoff value for reliability, Burns (1985) suggests a reliability coefficient of .70 to .90 is desirable and "considered to be sound statistical evidence of substantial reliability" (p.28).

Estimates of the stability and homogeneity of the instrument will be determined in this study. The reliability estimate for the stability of an instrument is represented by the correlation between two sets of scores resulting from administering an instrument on two separate occasions. Ideally the length of time between administrations should be no longer than two weeks (Impara & Stoker, 1985). This type of reliability is affected by how much the subjects are able to recall their answers from the previous administration.

Homogeneity refers to the internal consistency of an instrument. The procedures used for estimating this type of reliability include a correlation between scores on two halves of the instrument. Since the coefficient is based on only half of the instrument, the Spearman-Brown prophecy formula can be applied to elevate the estimate. Coefficient alpha and KR-20 are two other procedures, which, according to Impara and Stoker (1985), are similar in that they "represent estimates of the correlation between the measure under consideration and a hypothetical alternate (randomly parallel) form," (p. 69). KR-20 is applicable only

for tests which are scored dichotomously and represent an estimate of all possible split halves (Stanley & Hopkins, 1972). Coefficient alpha is more appropriate for attitude scales which are not being used in this study.

As mentioned in the review of literature, item analysis is a useful method for evaluating the quality of each test item and determining which items are in need of revision or elimination. Item analysis, as recommended by most measurement experts, should represent a comparison of those who scored in the upper 27% and the lower 27% of the total group. According to Payne (1974), this procedure keeps "the criterion group distinct enough to insure that they represent different levels of ability" (p. 275). Difficulty level, discrimination index, item response pattern are three procedures used to evaluate test items. Difficulty level is the proportion of correct responses from a group. A high difficulty index denotes an easy item in that a large number of individuals answered the item correctly. The difficulty level ranges from 0 to +1.00. Ebel (1972) suggests that items within the middle difficulty range of .25 to .75 contribute more to test reliability than items which are too easy or too hard.

Discrimination refers to the ability of an item to distinguish between high and low scorers. This ability is reflected in an index which is computed by "subtracting the number of students in the lower group who got the item right from the number in the upper group who got the item right and dividing by the number of students in either group" (Mehrens and Lehmann, p. 192, 1984). The index can range from -1.00 to

+1.00. A negative value indicates the need to revise or eliminate an item because more students in the low group answered the item correctly than in the high group. Mehrens and Lehmann (1984) suggest that it is not necessary for tests to include only items with high positive indices because low positive discrimination indices also contribute to the measurement of student abilities. Even though a low index indicates that the item was too easy, every test should include easy items which can serve as motivators. Also, removal of items which are too easy or too hard may affect the validity of the test. Ebel (1972) insists, however, that items with low or negative discriminatory ability should be revised. Ebel also discusses the influence of the mean discrimination index on reliability and determined that a 100 item test should have a mean index of at least .30 in order to achieve a reliability estimate above .70. Although discrimination ability and difficulty level are important, Mehrens and Lehmann (1984) note that "judgment of the quality of an item should be predicated more by whether or not it measures an important instructional objective rather than by the magnitude of its difficulty and discrimination indices" (p. 197).

Item response patterns are also an indication of item quality. Lindeman (1967) suggests that the effectiveness of item distractors is measured by the ability of the distractors to attract more students in the low group than in the high group. Often the cause of low discriminatory power is weak distractors. All distractors, according to Lindeman, should be used by some students; therefore, distractors which

have been ignored by all students should be revised to improve item quality.

In summary, issues related to value-added assessment, student involvement, and instrument quality based on its validity, reliability and item quality are all pertinent to the current study. Student scores on the Clothing & Textiles Core Knowledge Exam in its consolidated form and on the course specific components (subtests), quality credit average (QCA), final core course examination grades, and final core course grades are variables which will be used in this study to demonstrate criterion validity and construct validity, using correlational analyses. Interviews with faculty and students will demonstrate the content validity of the instrument. Characteristics of student involvement that will be considered are involvement in paid employment, involvement in student organizations, and involvement in academics. The procedures suggested for conducting value-added type research will also be considered in this study.

Hypotheses

The purpose of this study was to determine the validity, reliability, and item quality of a program-specific, university-level, faculty-developed assessment instrument designed to assess student mastery of core knowledge in the discipline. Additionally, the purpose was to determine the effectiveness of the instrument for ascertaining student outcomes from the value-added perspective and to determine whether performance was affected by student characteristics. The

instrument, referred to as the Clothing and Textiles Core Knowledge Exam, was evaluated for internal consistency, stability, content validity, construct validity, and criterion-related validity, item difficulty index, item discrimination index, and item distractor effectiveness. The following were objectives for this study:

1. To determine whether pretesting will unfairly influence the posttest scores of the pretest/posttest group.
2. To determine the predictive validity of the exam.
3. To determine the concurrent validity of the exam.
4. To determine the construct validity of the exam.
5. To determine the value-added quality of the exam.
6. To evaluate the exam for internal consistency, stability, and item quality.
7. To ascertain the content validity of the exam.
8. To determine whether characteristics of student involvement will affect outcomes on various measures of student achievement for the subtest sample and the upperclass sample.

Based on these objectives the following hypotheses were formulated:

Pretest Influence

Hypothesis 1: There will be no difference between posttest scores on the course specific components of the core knowledge exam for the pretest/posttest group and the posttest only group.

Criterion-related Validity

- Hypothesis 2: Pretest scores on the course-specific components (subtests) of the core knowledge exam will not be associated with final course grades.
- Hypothesis 3: Posttest scores on the course-specific components of the core knowledge exam will be positively associated with final course grades.

Construct Validity and Value-added Quality

- Hypothesis 4: Posttest scores on the course-specific components of the core knowledge exam will be positively associated with scores on the comprehensive final examinations in each course.
- Hypothesis 5: There will be no difference between pretest scores and posttest scores on the course-specific components of the core knowledge exam for the pretest/posttest group.
- Hypothesis 6: Scores on the core knowledge exam in its consolidated form will be positively associated with the quality credit average (QCA) for major courses for upperclassmen.
- Hypothesis 7: Upperclass CT majors will score higher on the core knowledge exam in its consolidated form than underclass CT majors.

Content Validity

- Hypothesis 8: The content of the Clothing and Textiles Core Knowledge Exam will be related to the course objectives and information taught in each of the five clothing & textiles core courses.

Reliability

- Hypothesis 9a: The internal consistency estimate for reliability for the subtests and for the core knowledge exam in its consolidated form will be at least .70.
- Hypothesis 9b: The stability estimate for reliability for the core knowledge exam in its consolidated form will be at least .70.

Item Quality

Hypothesis 10: The item analysis will yield information related to the reliability of the subtests and the full exam. The item analysis will relate to information gained from interviews with students and course instructors.

Student Involvement

Hypothesis 11: Student outcomes will be negatively affected by involvement in student organizations.

Hypothesis 12: Student outcomes will be negatively affected by involvement in paid employment.

Hypothesis 13: Student outcomes will be positively affected by the number of hours involved in school-related activities.

Assumptions and Limitations

The assumptions for this study were:

1. That students were sincere in their responses to the exam items and to the survey items.
2. That the difference between the pretest and posttest scores of the group taking both the pretest and the posttest is attributed to actual learning and not recollection of the pretest content.
3. That faculty were sincere in their efforts in developing questions for the exam.

Limitations

1. The study was limited to students currently enrolled in the Department of Clothing and Textiles and their performance on the core knowledge exam in its consolidated form and on the course specific components of the core knowledge exam.

2. The study was also limited to faculty experience with and knowledge of educational assessment.

Delimitation

A delimitation was that the study was limited to student knowledge acquired in the following courses: Clothing and People, Basic Apparel Assembly, Basic Apparel Assembly Laboratory, Introduction to the Fashion Industry, and Textile Science.

Operational Definitions

Clothing and Textiles Core Knowledge Exam -

an instrument which measures student knowledge related to course objectives covered in five core courses taught in the Department of Clothing and Textiles.

Course-specific components -

the 20 item subtests or portions of the core knowledge that relate to each of the five core courses in the Department of Clothing and Textiles.

Consolidated exam -

the core knowledge exam in its full 100 item form.

Construct Validity -

the ability of the Clothing and Textiles Core Knowledge Exam to (1) distinguish between knowledge acquired by upperclassmen and underclassmen, (2) distinguish between knowledge at the beginning of the course and at the end of the course, (3) yield student outcomes similar to the comprehensive final examination in each

course, and (4) yield student outcomes similar to the QCA in major for upperclassmen.

Content Validity -

the ability of the exam items to represent instructional objectives and information taught in the five core courses.

Criterion-related Validity -

the ability of the exam when given as a pretest and as a posttest to predict the students' final core course grades.

Reliability -

the internal consistency of the items on the core knowledge exam and the ability of the exam to yield consistent results over repeated trials.

Measures of Achievement -

instruments or indicators which reveal what the student has learned.

Value-added quality -

the ability of the Clothing and Textiles Core Knowledge Exam to detect educational gain from pretest to posttest and the ability of the exam to detect differences between those with little subject matter knowledge and those with a lot of subject matter knowledge.

CHAPTER III

METHODOLOGY

The purpose of this study was to determine the validity and reliability of a program-specific, university-level, faculty-developed assessment instrument designed to assess student mastery of core knowledge in the discipline. Additionally the purpose was to determine the effectiveness of the instrument for ascertaining student outcomes from the value-added perspective and to determine whether performance was affected by student characteristics. The instrument, referred to as the core knowledge exam, was evaluated for internal consistency, stability, content validity, construct validity, criterion-related validity, item difficulty level, item discriminatory power, and item distractor effectiveness. The collection and analysis of the data were conducted during the 1990-91 academic year in the Department of Clothing and Textiles at Virginia Tech. The methodology chapter will include (1) a description of the Department of Clothing and Textiles, (2) a description of the core courses, (3) the development of the instrument, (4) sampling procedures and collection of the data, and (5) the data analysis.

Description of the Department of Clothing and Textiles

The Department of Clothing and Textiles is one of five departments in the College of Human Resources at Virginia Tech. The department's

undergraduate curriculum offers concentrations in four option areas: Apparel Design, Fashion Merchandising, Fashion Merchandising and Marketing Education, and Textile Science. There are eleven faculty and approximately 158 students in the department. The department offers a total of twenty-six undergraduate courses. The core curriculum includes the five core courses around which the achievement test was designed. The five core courses and their course numbers are CT 1004 Clothing and People, CT 2404 Basic Apparel Assembly, CT 2414 Basic Apparel Assembly Laboratory, CT 2604 Introduction to the Fashion Industry, and CT 3204 Textile Science. These courses will be identified throughout the paper by course number.

Description of Core Courses

All courses are three semester credits except CT 2414 which is two semester credits. CT 1004 is a lecture course that provides an overview of the study of clothing from a behavioral and consumer perspective. CT 2404 is also a lecture course and concentrates on the apparel assembly process and apparel product quality. CT 2414 is a laboratory course which involves application of the principles and concepts taught in CT 2404. Students are required to construct garments according to specific guidelines. CT 2604 is a lecture course which focuses on various segments of the fashion industry as well as principles related to fashion. CT 3204 is a lecture and laboratory course involving the study of the chemical and physical properties of textile fibers and fabrics with an emphasis on consumer use.

Development of Instrument

The instrument which was used in this study is referred to as the Clothing and Textiles Core Knowledge Exam. The exam was developed in order to fulfill one of the initiatives in the Clothing and Textiles Department's five-year plan for assessing student outcomes. The exam was intended to be used as an outcomes measure to assess the student's knowledge of information covered in courses in the clothing and textiles core curriculum.

During the fall semester of 1988, instructors of the core courses were asked to submit questions representing their instructional objectives; these questions would be used in the development of a core examination. The compilation of questions was distributed to eleven faculty members in the department who were asked to analyze the content validity of the questions by matching the content with course objectives. Three faculty members responded to the request, and their suggestions were used to make revisions to the exam.

To pretest the instrument, the examination was administered to a group of junior clothing and textiles majors during the spring of 1989. Revisions made to the exam after the first pretest included adding and deleting items from each section to balance the test so each section contained twenty items. Items were also revised based on an item analysis indicating that certain items were not discriminating effectively.

During the spring of 1990 the revised examination was administered to a group of majors that included juniors and seniors. Faculty were

again asked to make revisions they felt necessary. Four items were changed during this revision.

Description of the Final Instrument

The final instrument consists of 100 multiple choice items relating to the five previously described core courses taught in the Department of Clothing and Textiles at Virginia Tech. For portions of the research, the final instrument was divided into five subtests representing each of the core courses (course-specific components). The subtests each contained twenty multiple choice items. On the first page of the pretest (subtest), posttest (subtest), and the consolidated exam Questions 1 through 9 requested demographic data. A copy of the complete exam and the demographic survey are in Appendix A. The key to each question was verified by consulting textbooks in Clothing and Textiles. Verification of the key is in Appendix F. Questions were related to student classification, option in clothing and textiles, and name of the five core courses taken or currently taking. For the pretest, other questions included name of major, reason for taking the core course in which the exam was administered, grade option for the course, and predicted course grade. For the posttest and the consolidated exam additional demographic data included status as a transfer student, number of out-of-class organizations involved in, number of hours per week spent on school related activities, involvement in paid employment for ten or more hours per week, and major-related employment.

Sampling Procedures and Data Collection

The sample for the subtests included students enrolled in the five previously described core courses (CT 1004, CT 2404, CT 2414, CT 2604, and CT 3204) enrolled during the fall semester 1990. The sample for the consolidated exam included upperclass majors (juniors and seniors) enrolled in the Economics of the Textile and Apparel Industry course, and underclass majors (freshmen and sophomores) in the department. Faculty who taught the five core courses comprised part of the sample for content validity.

The data were collected in three phases: (1) the course specific components of the exam (subtests) were administered in the five core courses; (2) the exam in its consolidated form (100 items) was administered to selected upperclass majors who had completed at least four of the core courses and to selected underclass majors who had taken fewer than three of the core courses; and (3) interviews were conducted during the subsequent semester with students from the five core courses who had taken the subtests, seniors who had taken the exam in its consolidated form, and faculty who taught the five core courses.

The process used to administer the five subtests was based on the separate-sample pretest-posttest quasi-experimental design described by Campbell and Stanley (1963). To implement this design, classes were tested with appropriate subtests on the first day of classes and on the last day of classes of the fall semester. The size of the sample, therefore, was dependent on the number of students attending classes when the data were collected. This design entails part of the sample

receiving a test prior to treatment (pretest) and the entire sample receiving a test after treatment (posttest). The entire sample receives the treatment, which, in this case, was instruction. This design allows for the determination of the following threats to internal validity: testing, regression, selection, and mortality. It also allows for the determination of the following threats to external validity: interaction of testing and treatment and interaction of selection and treatment. This design was followed as closely as possible.

In order to maintain as normal a classroom setting as possible, the instructor of each of the core courses administered the course specific components (subtests) of the core knowledge exam on the first day of classes for the semester. The actual subtest and a dummy test were stacked in an alternate sequence in order to obtain a somewhat random sample of students during the pretest phase of the research. Each student had equal opportunity of receiving the actual subtest. One instructor removed the dummy tests from the stack and administered the actual subtest to all students. Therefore, for the Textile Science course all students attending class on the first day received the course specific component of the core knowledge exam.

The students used opscan forms to mark their answers to the subtest items. All students were assured anonymity; however, student identification numbers were requested so the pretest scores could later be matched with the posttest scores and student grades. The subtests were administered as a posttest to all students attending class on the last day of the fall semester 1990 after instruction and before the

final examination in each of the five core courses. Final course grades and examination scores were obtained for all of the students from course grade reports.

The core knowledge exam in its consolidated form was administered to upperclass and underclass clothing and textiles majors during the fall semester 1990. For the test-retest reliability, the consolidated exam was administered to a group of upperclassmen who were enrolled in a senior level course in the spring of 1991. The testing periods were one week apart.

The faculty teaching the core courses were interviewed regarding student attendance throughout the semester and the correlation between the course content and the content of the subtest. Although student identity was protected, faculty identity could not be protected as it is possible to identify faculty who taught the courses used in the study. A group of students from each of the five core courses were selected based on their performance on the posttest (high achievers and low achievers) and interviewed regarding the amount of time they spent studying for the course, their attendance in the course, and their opinion regarding the correlation between the items on the core knowledge exam and information covered in class. The students were also asked their opinion regarding the relationship between content in the core courses and other courses in the department and outside the department. Ten upperclass majors who took the exam in its consolidated form were also interviewed. For an accurate reporting of objectives and subtest items which the students felt were emphasized during the

semester, at least three students in CT 1004, CT 2404, CT 2414, CT 2604, two students in CT 3204, and four seniors must be in agreement. The protocol for the interviews with faculty and students is in Appendix A. All interviews were conducted with individuals and not in groups. Each interview took less than twenty minutes to complete.

Analysis of Data

The data used in this study were of two types: quantitative and qualitative. The quantitative data consisted of the exam scores and grades received by the students. The qualitative data consisted of the interviews conducted with students and faculty. Only students who remained in the courses throughout the semester were included in the final sample and reported in this study.

Statistical tests used in this study for the quantitative data were descriptive statistics, correlational statistics, and statistical test of differences. Data from each of the five core courses were analyzed independently and collectively (across courses). Number Cruncher Statistical System (Hintze, 1987) was used to accomplish the data analysis.

Descriptive statistics were used to report frequencies, percentage distribution, and means for the subtest and consolidated exam results and for the demographic characteristics of the sample. Descriptive statistics were also used to report the item analysis for the core knowledge exam.

Correlational statistics used in the study included tests for internal consistency and stability. The Kuder-Richardson-20 formula and split-half reliability with Spearman Brown were used as reliability tests for internal consistency. Pearson product-moment correlation was used to compare two sets of exam scores to obtain an estimate of stability. Pearson product-moment correlation was also used to test for relationships between variables of interest and the course specific components of the core knowledge exam and the consolidated exam.

Criterion-related validity was ascertained by the correlation between the pretest scores and final course grades, and by the correlation between the posttest scores and final course grades. The construct which was investigated was knowledge of the core course subject-matter represented by the exam items. Construct validity, therefore, was demonstrated by the correlation between the posttest scores and the comprehensive final examination grades, the comparison of consolidated exam scores for upperclass majors and underclass majors, and the comparison of the pretest scores with the posttest scores. Content validity was demonstrated by comparing pretest and posttest scores of those taking both the pretest and the posttest.

Statistical tests of differences used in this study included the independent sample t-test and the matched sample t-test to compare scores on the core knowledge exams between groups. Analysis of variance was used to compare students on achievement outcomes as affected by involvement factors.

For the qualitative data a descriptive analysis of the interviews with faculty and students was conducted. The relationship between the interviews and the item analysis was also included in the descriptive analysis. Content validity was demonstrated by the responses from students and faculty regarding the correlation between items on the subtests and subject matter taught in the courses. A summary of analyses is presented in Table I.

Table 1. Summary of Data Analysis by Hypothesis

Hypothesis	Statistical Procedure	Variables
Hypothesis 1	Independent t-test	posttest scores of pretest group & posttest-only group
Hypothesis 2	Pearson product-moment correlation	pretest scores and course grade
Hypothesis 3	Pearson product-moment correlation	posttest scores and course grade
Hypothesis 4	Pearson product-moment correlation	posttest scores and final exam grade
Hypothesis 5	Matched sample t-test	pretest scores and posttest scores
Hypothesis 6	Pearson product-moment correlation	consolidated exam score, QCA in major
Hypothesis 7	Independent t-test	consolidated exam scores for upper-class majors & underclass majors
Hypothesis 8	(Descriptive analysis of interviews)	
Hypothesis 9	Pearson product-moment correlation Spearman-Brown Formula KR-20 coefficient	subtests & consolidated exam
Hypothesis 10	Difficulty level Discriminatory index	subtests & consolidated exam
Hypothesis 11	Independent t-test	number of organizations posttest scores, consolidated exam scores, QCA for upperclass majors, course grades, final exam grades
Hypothesis 12	Independent t-test	employment status, posttest scores, consolidated exam scores, QCA for upperclass majors, course grades, final exam grades
Hypothesis 13	One-way ANOVA	hours spent on school work, posttest scores, consolidated exam scores, QCA for upperclass majors course grades, final exam grades

CHAPTER IV

RESULTS AND DISCUSSION

The results of the study are presented in this chapter. Included in the results are a description of the sample, results of the data analysis by objective. The .05 level of significance was established for all statistical tests.

Description of the Sample

The sample for this study involves two major groups: the sample participating in the evaluation of the subtests (the course specific components of the core knowledge exam) and the sample participating in the evaluation of the core knowledge exam in its consolidated form.

Subtest Sample

Results have been reported for students who remained in the class throughout the semester and for whom subtest scores and final course grades were available. The total sample size was 129 individual students. There were thirty-five students who were enrolled in more than one of the courses and therefore participated more than once resulting in total student scores of 179. Frequencies by course for the sample participating in the evaluation of the subtests are presented in Table 2 in three groups: pretest-only, pretest/posttest, and posttest-only. As shown in Table 2 very few students were in the pretest-only group which represents students who did not attend class when the

Table 2. Frequencies of the Sample Participating in the Subtest Evaluations

Course	<u>Pretest Only</u>		<u>Pretest/Posttest</u>		<u>Posttest Only</u>		<u>Total</u>	
	N	%	N	%	N	%	N	%
CT 1004	10	15.6	20	31.3	34	53.1	64	100
CT 2404	0	0.0	19	47.5	21	52.5	40	100
CT 2414	3	9.1	14	42.4	16	48.5	33	100
CT 2604	3	10.3	10	34.5	16	55.2	29	100
CT 3204	1	7.69	11	84.6	1	7.69	13	100
Total	17		74		88		179**	

** Represents an overlap of 35 students

***No student took the same test twice

posttest was administered. Those in the posttest-only group were either absent on the first day of class or were administered the dummy test at the beginning of the semester.

The demographic characteristics of this sample by course are presented in Tables 3, 3a, and 3b. The classification of the subjects for each course were representative of the course levels, except CT 1004 which, although a freshmen level course, is taken as an elective by many students outside of the clothing and textiles department. The student classification in CT 1004 was extremely varied consisting mostly of juniors (35.9%), and non majors (65.6%).

Variables in Table 4 pertain to questions asked only on the posttest, which is the reason the totals are different from those in Table 3. Many students, as shown in the table, were involved in extra-curricular organizations. For most (73%) of the students, the number of organizations was fewer than three. In terms of schoolwork, the majority of the sample in each class spent fewer than thirty hours per week in that capacity. Part of the reason could be attributed to their involvement in organizations or their involvement in paid employment, in which case approximately thirty percent of the whole sample worked. For most subjects, the type of employment was not related to their major.

Consolidated Exam Sample

The sample receiving the core knowledge exam in its consolidated form (100 items) consisted of twenty-four underclassmen and twenty-four upperclassmen. The demographic characteristics of this sample are presented in Table 5. The majority (21) of the underclassmen were

Table 3. Demographic Characteristics of the Total Subtest Samples

Characteristic	CT 1004		CT 2404		CT 2414		CT 2604		CT 3204	
	N	%	N	%	N	%	N	%	N	%
<u>Classification</u>										
Freshman	14	21.9	9	22.5	6	18.2	5	17.2	0	0
Sophomore	15	23.4	21	52.5	18	54.5	18	62.1	0	0
Junior	23	35.9	9	22.5	9	27.3	5	17.2	9	69.2
Senior	12	18.8	1	2.5	0	0.0	1	3.4	4	30.8
Total	64	100.0	40	100.0	33	100.0	29	100.0	13	100.0
<u>CT Major</u>										
Yes	22	34.4	38	95.0	31	93.9	22	75.9	13	100.0
No	42	65.6	2	5.0	2	6.5	7	24.1	0	0.0
Total	64	100.0	39	100.0	33	100.0	29	100.0	13	100.0

Table 3a. Demographic Characteristics of the Pretest/Posttest Group

Characteristic	CT 1004		CT 2404		CT 2414		CT 2604		CT 3204	
	N	%	N	%	N	%	N	%	N	%
<u>Classification</u>										
Freshman	6	30.0	4	21.1	3	21.4	1	10.0	0	0.0
Sophomore	3	15.0	12	63.2	9	64.3	7	70.0	0	0.0
Junior	7	35.0	3	15.8	2	14.3	1	10.2	9	81.8
Senior	4	20.0	0	0.0	0	0.0	1	10.0	2	18.2
Total	20	100.0	19	100.0	14	100.0	10	100.0	11	100.0
<u>CT Major</u>										
Yes	7	35.0	18	94.7	13	92.9	10	100.0	11	100.0
No	13	65.0	1	5.3	1	7.1	0	0.0	0	0.0
Total	20	100.0	19	100.0	14	100.0	10	100.0	11	100.0

Table 3b. Demographic of Pretest-Only and Posttest Only Groups

Characteristic	CT 1004		CT 2404		CT 2414		CT 2604		CT 3204	
	N	%	N	%	N	%	N	%	N	%
<u>Classification</u>										
Freshman	8	18.2	5	23.8	3	15.8	4	21.0	0	0.0
Sophomore	12	27.3	9	42.8	9	47.4	11	58.9	0	0.0
Junior	16	36.3	6	28.6	7	36.8	4	21.0	0	0.0
Senior	8	18.2	1	4.8	0	0.0	0	0.0	2	100.0
Total	44	100.0	21	100.0	19	100.0	19	100.0	2	100.0
<u>CT Major</u>										
Yes	15	34.1	20	95.2	18	94.7	12	63.1	2	100.0
No	29	65.9	1	4.8	1	5.3	7	36.8	0	0.0
Total	44	100.0	21	100.0	19	100.0	19	100.0	2	100.0

Table 4. Additional Characteristics of the Total Posttest Sample

Characteristic	CT 1004		CT 2404		CT 2414		CT 2604		CT 3204	
	N	%	N	%	N	%	N	%	N	%
<u>Number of Organizations</u>										
Fewer than 3	31	57.4	35	87.5	25	83.3	20	76.9	7	58.3
3 or more	15	27.8	4	10.0	4	13.3	6	23.1	5	41.7
No response	8	14.8	1	2.5	1	3.3	0	0.0	0	0.0
Total	54	100.0	40	100.0	30	100.0	26	100.0	12	100.0
<u>Hours Spent on School Work</u>										
20 or Less	23	42.6	15	37.5	13	43.3	8	30.8	4	33.3
21 - 29	14	25.9	13	32.5	9	30.0	8	30.8	8	66.7
30 - 39	9	16.7	6	15.0	5	16.7	6	23.1	0	0.0
40 or more	2	3.7	5	12.5	2	6.7	4	15.3	0	0.0
No Response	6	11.1	1	2.5	1	3.3	0	0.0	0	0.0
Total	54	100.0	40	100.0	30	100.0	26	100.0	12	100.0
<u>Involvement in Paid Employment</u>										
Yes	15	27.8	12	30.0	13	43.3	7	26.9	5	41.7
No	33	61.1	27	67.5	16	53.3	19	73.1	7	58.3
No Response	6	11.1	1	2.5	1	3.3	0	0.0	0	0.0
Total	54	100.0	40	100.0	30	100.0	26	100.0	12	100.0
<u>Employment Related to Major</u>										
Yes	7	13.0	5	12.5	7	23.3	2	7.7	2	16.7
No	10	18.5	10	25.0	6	20.0	5	19.2	4	33.3
Doesn't Apply	31	57.4	24	60.0	16	53.3	19	73.1	6	50.0
No Response	6	11.1	1	2.5	1	3.3	0	0.0	0	0.0
Total	54	100.0	40	100.0	30	100.0	26	100.0	12	100.0

Table 4a. Additional Characteristics of the Pretest/Posttest Group

Characteristic	CT 1004		CT 2404		CT 2414		CT 2604		CT 3204	
	N	%	N	%	N	%	N	%	N	%
<u>Number of Organizations</u>										
Less than 3	13	65.0	16	84.2	13	92.9	8	80.0	6	54.5
3 or more	7	35.0	3	15.8	1	7.1	2	20.0	5	45.5
Total	20	100.0	19	100.0	14	100.0	10	100.0	11	100.0
<u>Hours Spent on School Work</u>										
20 or Less	8	40.0	7	36.8	4	28.6	3	30.0	4	36.4
21 - 29	4	20.0	4	21.1	5	35.7	4	40.0	7	63.6
30 - 39	5	25.0	4	21.1	3	21.4	2	20.0	0	0.0
40 or more	1	5.0	3	15.8	1	7.1	1	10.0	0	0.0
No Response	2	10.0	1	5.2	1	7.1	0	0.0	0	0.0
Total	20	100.0	19	100.0	14	100.0	10	100.0	11	100.0
<u>Involvement in Paid Employment</u>										
Yes	4	20.0	4	21.0	7	50.0	1	10.0	5	45.5
No	14	70.0	14	73.7	6	42.9	9	90.0	6	54.5
No Response	2	10.0	1	5.3	1	7.1	0	0.0	0	0.0
Total	20	100.0	19	100.0	14	100.0	10	100.0	11	100.0
<u>Employment Related to Major</u>										
Yes	3	15.0	2	10.5	2	14.3	1	10.0	2	18.2
No	2	10.0	4	21.0	5	35.7	0	0.0	4	36.4
Doesn't Apply	13	65.0	12	63.2	6	42.9	9	90.0	5	45.4
No Response	2	10.0	1	5.3	1	7.1	0	0.0	0	0.0
Total	20	100.0	19	100.0	14	100.0	10	100.0	11	100.0

Table 4b. Additional Characteristics of the Posttest-Only Group

Characteristic	CT 1004		CT 2404		CT 2414		CT 2604		CT 3204	
	N	%	N	%	N	%	N	%	N	%
<u>Number of Organizations</u>										
Less than 3	26	76.5	20	95.2	13	81.3	12	75.0	1	100.0
3 or more	8	23.5	1	4.8	3	18.8	4	25.0	0	0.0
Total	34	100.0	21	100.0	16	100.0	16	100.0	1	100.0
<u>Hours Spent on School Work</u>										
20 or Less	15	44.1	8	38.1	9	56.2	5	31.2	0	0.0
21 - 29	10	29.4	9	42.9	4	25.0	4	25.0	1	100.0
30 - 39	4	11.8	2	9.5	2	12.5	4	25.0	0	0.0
40 or more	1	2.9	2	9.5	1	6.3	3	18.8	0	0.0
No Response	4	11.8	0	0.0	0	0.0	0	0.0	0	0.0
Total	34	100.0	21	100.0	16	100.0	16	100.0	1	100.0
<u>Involvement in Paid Employment</u>										
Yes	11	32.3	48	38.1	6	37.5	6	37.5	0	0.0
No	19	55.9	13	61.9	10	62.5	10	62.5	1	100.0
No Response	4	11.8	0	0.0	0	0.0	0	0.0	0	0.0
Total	34	100.0	21	100.0	16	100.0	16	100.0	1	100.0
<u>Employment Related to Major</u>										
Yes	4	11.8	3	14.3	5	31.3	1	6.2	0	0.0
No	8	23.5	6	28.6	1	6.2	5	31.3	0	0.0
Doesn't Apply	18	52.9	12	57.1	10	62.5	10	62.5	1	100.0
No Response	4	11.8	0	0.0	0	0.0	0	0.0	0	0.0
Total	34	100.0	21	100.0	16	100.0	16	100.0	1	100.0

Table 5. Demographic Characteristics of Sample Taking the Consolidated Exam

Characteristic	<u>Upperclassmen</u>		<u>Underclassmen</u>	
	N	%	N	%
<u>Student Classification</u>				
Freshmen	0	0.0	21	87.5
Sophomore	0	0.0	3	12.5
Junior	1	4.2	0	0.0
Senior	23	95.8	0	0.0
Total	24	100.0	24	100.0
<u>Clothing & Textiles Major</u>				
Yes	24	100.0	21	87.5
No	0	0.0	3	12.5
Total	24	100.0	24	100.0
<u>Course-taking Pattern</u>				
CT 1004				
Have Not Taken	0	0.0	13	54.2
Now Taking	0	0.0	6	25.0
Have Taken	24	100.0	5	20.8
Total	24	100.0	24	100.0
CT 2404				
Have Not Taken	0	0.0	24	100.0
Now Taking	0	0.0	0	0.0
Have Taken	24	100.0	0	0.0
Total	24	100.0	24	100.0
CT 2414				
Have Not Taken	0	0.0	23	95.8
Now Taking	0	0.0	1	4.2
Have Taken	24	100.0	0	0.0
Total	24	100.0	24	100.0
CT 2604				
Have Not Taken	0	0.0	14	58.3
Now Taking	1	4.2	10	41.7
Have Taken	23	95.8	0	0.0
Total	24	100.0	24	100.0
CT 3204				
Have Not Taken	3	12.5	24	100.0
Now Taking	3	12.5	0	0.0
Have Taken	18	75.0	0	0.0
Total	24	100.0	24	100.0

Table 5. Demographic Characteristics of Sample Taking the Consolidated Exam (Continued)

Characteristic	<u>Upperclassmen</u>		<u>Underclassmen</u>	
	N	%	N	%
<u>Number of Out-of-Class Organizations</u>				
Fewer than 3	12	50.0	17	70.8
3 or more	12	50.0	4	16.7
No Response	0	0.0	3	12.5
Total	24	100.0	24	100.0
<u>Hours Spent on Schoolwork</u>				
20 or less	0	0.0	1	4.2
21 - 29	9	37.5	8	33.3
30 - 39	9	37.5	8	33.3
40 or more	5	20.8	5	20.8
No Response	1	4.2	2	8.3
Total	24	100.0	24	100.0
<u>Involvement in Paid Employment</u>				
Yes	16	66.7	1	4.2
No	8	33.3	22	91.6
No Response	0	0.0	1	4.2
Total	24	100.0	24	100.0
<u>Employment Related to Major</u>				
Yes	6	25.0	0	0.0
No	9	37.5	1	4.2
No Response	9	37.5	23	95.8
Total	24	100.0	24	100.0

freshmen and the majority (24) of the upperclassmen were seniors. All were majors in the Department of Clothing and Textiles except three of the underclassmen.

In terms of core course exposure in the core curriculum, the majority (17) of the upperclassmen had taken all five core courses. The remainder (7) had taken at least four of the courses. The underclassmen, as expected and desired, had not taken most of the core courses, especially CT 3204 and CT 2604. Eight students had taken none of the courses, eleven students were concurrently enrolled in one or two of the courses and had not taken any others, and five students had taken one of the courses and were concurrently enrolled in one of the other courses.

Half (12) of the upperclassmen were involved in fewer than three student organizations as compared to 17 of the underclassmen. For the majority of both groups, time spent on schoolwork ranged from twenty-one to thirty-nine hours per week. Most of the seniors (16) and only one underclassman were involved in paid employment. In most cases, the employment was not related to their major.

Subtest and Exam Results

The mean, standard deviation, and range of scores for the pretest, posttest, and consolidated exam are presented in Table 6. The pretest mean for CT 2414 (35.88%) was the lowest of the five courses, and the posttest mean (70%) for the same course was the highest of the five courses, suggesting that the value-added was greater for this class than for the other four courses. The course which appeared to have the least value-added was CT 2404; it had a pretest mean of 49.21 and a posttest

Table 6. Results of Subtests and Consolidated Exam

Test	Pretest				Posttest			
	n	<u>M</u>	<u>SD</u>	Range	n	<u>M</u>	<u>SD</u>	Range
<u>Subtests</u>								
CT 1004	30	44.17	13.59	20-75	54	63.98	14.77	20-95
CT 2404	19	49.21	7.31	35-65	40	55.63	12.62	20-75
CT 2414	17	35.88	18.31	10-85	30	70.00	12.39	35-95
CT 2604	13	50.00	13.39	30-70	26	67.69	10.51	45-90
CT 3204	12	36.67	9.37	20-50	12	65.00	16.79	30-80
Across Courses	91	43.52	13.95	10-85	162	63.70	14.24	20-95
<u>Consolidated Exam</u>								
Upperclassmen					24	68.67	9.70	54-88
Underclassmen					24	45.92	10.56	26-72

mean of 55.63. Across courses on the pretest, student scores ranged from 10 to 85. On the posttest, student scores ranged from 20 to 95, illustrating an increase in performance.

The overall standard deviation was 13.95 on the pretest and 14.24 on the posttest, indicating that the scores were widely dispersed. On the consolidated exam, the mean score for upperclassmen (68.87) was higher than that for underclassmen (45.92), with similar standard deviations.

Pretest Influence

Hypothesis 1: There will be no difference between posttest scores on the course specific components of the core knowledge exam for the pretest/posttest group and the posttest only group.

A t-test analysis was conducted for all but one course (CT 3204) which did not have a substantial sample size in the posttest only groups for reasons mentioned earlier. As shown in Table 7 independent t-tests for each course and across courses were not significant at the .05 level, thus leading to the acceptance of this hypothesis. Since the sample sizes for the pretest/posttest and the posttest-only groups were not equivalent, an F-test for the homogeneity of variance was conducted. The F ratio for each course and across courses was not significant at the .05 level.

The results of these analyses reveal that those in the pretest/posttest groups, who were exposed to the test at the beginning of the semester did not have an advantage over those who were only

Table 7. Comparison of Posttest Scores of the Pretest/Posttest Group and the Posttest-Only Group

Course	<u>Pretest/Posttest Group</u>			<u>Posttest Only Group</u>			t-value
	<u>n</u>	<u>M</u>	<u>SD</u>	<u>n</u>	<u>M</u>	<u>SD</u>	
CT 1004	20	66.75	15.58	34	63.25	14.26	1.06
CT 2404	19	57.63	13.88	21	53.81	12.03	.9328
CT 2414	14	70.00	14.94	16	70.00	10.16	0
CT 2604	10	70.00	9.43	16	66.25	11.18	.8811
Across Courses	74	65.68	14.48	88	62.04	13.89	-1.62

tested once. Although the means for the pretest/posttest group are higher in all but one course, the differences were not significant.

Criterion-related Validity

Hypothesis 2: Pretest scores on the course specific components (subtests) of the core knowledge exam will not be associated with final course grades.

The Pearson product-moment correlations between pretest scores and final course grades for each of the core courses were low and not significant at the .05 level of significance (Table 8). The correlation across courses was also low (.0969) and not significant, indicating that pretest scores were not associated with final grades and leading to the acceptance of this hypothesis. The purpose of this hypothesis was to determine whether the core knowledge exam, when given as a pretest, could predict success in the applicable courses. These results suggest that the core knowledge exam is not an effective predictor of student performance when given as a pretest. According to Schwarz (1971), tests used for prediction contain items which measure past achievements, not expected achievements as the Clothing and Textiles Core Knowledge Exam does.

Hypothesis 3: Posttest scores on the course specific components of the core knowledge exam will be positively associated with final course grades.

To the extent that the subtest portions of the core knowledge exam represent information instructors feel is necessary to succeed in the

Table 8. Product-Moment Correlations Between Course Grades and Pretest Scores and Posttest Scores by Course and Over all Courses

Course	Coefficient			
	Pre-test to Grade	p-value	Posttest to Grade	p-value
CT 1004	-.0891	.7089	.2396	.0810
CT 2404	.0011	.9965	.2140	.1848
CT 2414	.1221	.6777	.1073	.5724
CT 2604	-.1899	.5993	.2684	.1850
CT 3204	-.3061	.3599	.6400*	.0250
Across courses	.0969	.3635	.1960*	.0124

*Statistical value meets pre-established significance level

courses, the relationship between the exam and course grades should be high and positive. The Pearson product-moment correlations between posttest scores and final course grades were low and positive and not significant for all courses except CT 3204 (Table 8). The Pearson product-moment correlation coefficient (.6400) between posttest scores and course grades in CT 3204 was statistically significant at $p=.0250$. Overall, for all courses, the correlation (.1960) between posttest scores and course grades, although low, was also significant at $p=.0124$, indicating that the subtest, when given as a posttest, may predict final course grades in CT 3204.

These results suggest a direct relationship between scores on the core knowledge exam after students have received instruction and final course grades. Therefore the hypothesis was accepted. This relationship demonstrates to some extent the concurrent validity of the exam, that it does measure degree of student success in the core courses.

Construct Validity

Hypothesis 4: Posttest scores on the course specific components of the core knowledge exam will be positively associated with scores on the comprehensive final examinations in each course.

Final examinations for all but one of the core courses (CT 1004) were comprehensive. Consequently, analyses were conducted for all courses except CT 1004. The Pearson product-moment correlation (.3249) between posttest scores and final examination scores in CT 2404 was

significant at $p=.0408$ (Table 9). Low positive correlations were not significant for the other three core courses. For the data across courses, the correlation (.4844) was significant at $p=.0002$; therefore, the hypothesis was accepted. The R^2 value was .2347, indicating that the two variables shared only 23% of the variance. These results are not surprising considering the brevity of the subtests (20 items) as compared to the length of the final examination. Some evidence of construct validity, however, was demonstrated by these data.

Hypothesis 5: Posttest scores for the pretest/posttest group will be higher than pretest scores on the course specific components of the core knowledge exam.

A paired t-test was conducted to test the relationship between pretest scores and posttest scores for the pretest/posttest group. As shown in the Table 10, the differences between the two scores were significant for all courses and across courses leading to the acceptance of this hypothesis. There was a significant increase in the exam score from pretest to posttest. These results illustrated the ability of the core knowledge exam to measure knowledge acquired in each of the core courses, thereby providing further evidence of construct validity and the ability of the exam to be used as a value-added measure. These results also provide some evidence of content validity (i.e., that there is some correlation between information taught in the course and the items on the exam).

Hypothesis 6: Scores on the core knowledge exam in its consolidated form will be positively associated with the quality credit average (QCA) for major courses for upperclassmen.

Table 9. Correlation Coefficients for Posttest Scores and Final Examination Grades

Course	Coefficient	p-values
CT 2404	.3249*	.0408
CT 2414	.2311	.2191
CT 2604	.3693	.0634
CT 3204	.5021	.0962
Across Courses	.4844*	.0002

*Statistical value meets pre-established significance level.

Table 10. A Comparison of Pretest and Posttest Scores for the Pretest/Posttest Group

Course	n	<u>Pretest Score</u> <u>M</u>	<u>Posttest Score</u> <u>M</u>	<u>SE</u>	t-value
CT 1004	20	43.25	66.75	3.86	6.09***
CT 2404	19	49.21	57.63	3.29	2.56*
CT 2414	14	37.86	70.00	4.15	7.74***
CT 2604	10	52.00	70.00	4.54	3.95**
CT 3204	11	35.91	68.18	5.37	6.01***
Across Courses	74	43.78	65.68	2.09	10.50***

* $p \leq .05$

** $p \leq .01$

*** $p \leq .001$

Scores on the core knowledge exam ranged from 54 to 88 and the quality credit average (QCA) of upperclassmen ranged from 2.3 to 3.9 on a 4.0 scale. Exam scores were correlated with QCA in major courses for the upperclassmen who took the consolidated exam. The Pearson product-moment correlation (.8500) was high and positive and highly significant at $p = .0001$. Thus students who scored higher on the core knowledge exam also had higher QCAs in their major and the hypothesis was accepted. These results provide some evidence of construct validity to the extent that the exam measures the level of knowledge of clothing and textiles core information, and the QCA in major is an indication of the level of knowledge of information related to clothing and textiles.

Hypothesis 7: Upperclass CT majors will score higher on the core knowledge exam in its consolidated form than underclass CT majors.

Means and standard deviations of the two groups were computed; differences between means were analyzed by an independent t-test. As shown in Table 11, upperclassmen scored significantly higher on the core knowledge exam than underclassmen and the hypothesis was accepted. This result was expected since most of the seniors had taken at least four of the core courses and the underclassmen had taken no more than two. The results of this analysis demonstrate that the exam is a viable measure of knowledge acquired in the clothing and textiles curriculum, providing additional evidence of construct validity and the ability of the exam to measure value-added knowledge using a cross-sectional approach.

Table 11. Comparison of Core Knowledge Exam Scores Based on Student Classification

Group	<u>M</u>	<u>SD</u>	t-value
Upperclassmen	68.67	9.70	7.77*
Underclassmen	45.92	10.56	

*p < .01

Content Validity

Hypothesis 8: The content of the Clothing and Textiles Core Knowledge Exam will be related to the course objectives and information taught in each of the five clothing & textiles core courses.

To establish the content validity of the subtests and the core knowledge exam in its consolidated form, interviews were conducted during the 1991 spring semester with faculty who taught the core courses, with a subset of students from each of the five core courses who had taken the subtests, and with a subset of seniors in the department who had taken the full exam. The interview protocol for faculty and students is in Appendix B. The results of the interviews will be presented by course.

CT 1004 - Clothing and People

The instructor of this course and eight students were interviewed during individual sessions. The average score on the posttest for all students was 63.98. Four of the students interviewed had earned a score of 75 on the posttest; one student each had earned a score of 85, 80, 45, and 40.

The course outline for CT 1004 contained five objectives as presented in Table 12. Subtest items related to the objectives are in Appendix C. The instructor indicated that more emphasis was placed on objectives 2 and 4 than on the other three. The objectives were used as a guide to organize the course and to develop test questions. In terms of the items for the subtest, the instructor indicated that all items

Table 12. Course Objectives For CT 1004 - Clothing and People

-
1. Appreciate the relationships between clothing and textiles and the culture in which we live.
 2. Understand the importance of dress as a means of self expression and a mode of communication among individuals, groups, and cultures.
 3. Understand the forces that influence clothing behavior.
 4. Recognize the components of apparel-fibers, fabrics, and finishes-which help individuals make clothing choices that provide maximum satisfaction.
 5. Increase satisfaction in personal use of clothing.
-

were covered during the semester, but that eight of the twenty items were given the most emphasis.

All of the students confirmed that a syllabus was distributed and included the objectives in Table 12. Of the five objectives on the syllabus, at least three students agreed that objectives 2 and 4 were given the most emphasis, which agreed with the instructor. Three students also felt objectives 1 and 3 were emphasized. All but one of the students interviewed felt the learning objectives had been achieved.

In reference to items on the subtest, three or more students agreed that twelve items were given the most emphasis. Of the twelve items, six corresponded with those selected by the instructor. One of the six items related to the objective 2 which both groups felt were emphasized; none related to objective 4. The other five related to objectives the students recalled as being emphasized. The objective which neither the students nor the instructor felt had been emphasized had no subtest items related to it. This information will be helpful in partially explaining results of the item analysis.

CT 2404 - Basic Apparel Assembly Lecture

The instructor and eight students were interviewed in reference to this course. The average score for all students on the posttest for this course was 55.63. The students interviewed scored the following on the subtest: 75, 70, 60 (2), 55 (3), and 35.

The course outline for CT 2404 contained the four objectives listed in Table 13. Subtest items related to the objectives are presented in Appendix C. According to the instructor, more emphasis was placed on

objectives 1 and 4. The objectives were used primarily to organize the course and develop tests. After reviewing the items on the subtest, the instructor acknowledged that all items were covered during the semester; however, seven of the twenty were given the most emphasis. All of the students interviewed recalled getting a syllabus with the objectives from Table 13. Although all of the objectives were covered, there was not a total consensus on which objectives were given the most emphasis. At least three of the eight students felt more emphasis was placed on objectives 1 and 4, which concurred with the instructor's assessment. Five of the students felt they had achieved the four learning objectives.

In reference to the subtest related to CT 2404, at least three or more students agreed that seven of the twenty items were emphasized. Three of the seven items were the same ones identified by the instructor. These three items also related to the objectives that both the instructor and the students felt had been emphasized in the course.

CT 2414 - Basic Apparel Assembly Laboratory

The instructor of this course and seven students were interviewed regarding the content of this course. The average score based on all students taking the posttest for this course was 70. Student scores on the subtest portion of the core knowledge exam were 95, 85 (2), 80, 70, 60, and 35.

The instructor of this course was also the instructor for CT 1004. The course syllabus distributed on the first day of class included the objectives presented in Table 14. Test items related to the objectives

**Table 13. Course Objectives for CT 2404 - Basic Apparel Assembly
Lecture**

-
1. Interpret terminology associated with the manufacturing, merchandising, and consumer use of apparel products.
 2. Understand basic concepts and principles that govern production processes.
 3. Understand the importance of efficiently managing time, equipment, and supplies in the garment assembly process.
 4. Understand elements in the assessment of apparel quality.
-

Table 14. Course Objectives for CT 2414 - Basic Apparel Assembly Laboratory

-
1. Understand and apply terminology used in the apparel process.
 2. Utilize concepts, principles, and techniques of garment assembly in the production of finished consumer products.
 3. Demonstrate through garment construction the importance of time, equipment, supplies, and fabric in evaluating garment quality.
 4. Apply some basic fitting principles in the construction of garments.
-

are in Appendix C. According to the instructor, the objectives were used to organize the course and to serve as a guide in developing tests. The objectives, she reported, were all covered during the semester. The ones which were emphasized the most were objectives 1 and 3. In reference to the items on the subtest for CT 2414, the instructor identified eight items which were covered in the most detail.

When shown the course objectives in Table 14, all students agreed that the objectives were included on the syllabus. According to the students, all objectives were covered. At least three students agreed that objectives 1 and 2 were emphasized. All of the students said they had achieved each of the learning objectives to some extent.

When asked about the items on the subtest, two students felt all items were given equal attention. At least three other students felt nine items were emphasized, six of which also corresponded with the instructor's assessment. Four of the six items related to objective 1, which both the instructor and the students felt had been emphasized. The other two items related to objective 2, which only the students felt had been emphasized.

CT 2604 - Introduction to the Fashion Industry

The instructor and seven students were interviewed during individual sessions. The average score on the posttest based on all student scores in the course was 67.69. The students' scores on the posttest were 80, 75, 70 (2), 65 (2), 60.

The course syllabus for CT 2604 contained the five objectives presented in Table 15. Subtest items which relate to the objectives are

Table 15. Course Objectives for CT 2604 - Introduction to the Fashion Industry

-
1. To acquire a basic understanding of the nature of fashion.
 2. To analyze the economic, sociological, and psychological factors affecting contemporary fashion; to become acquainted with literature and resource materials dealing with these factors.
 3. To understand the history and development of the various segments of the fashion industry and the unique and complete interdependency of these segments.
 4. To gain an understanding and appreciation of the purposes served by foreign and domestic markets and to examine the role of designers in these markets.
 5. To acquire a realistic understanding of the opportunities and responsibilities in the fashion industry.
-

in Appendix C. The instructor recalled using the objectives to organize the course and to develop tests and lecture materials. In her opinion, all objectives were covered in class with greater emphasis placed on objectives 1, 2, 4, and 5.

In reference to the subtest items, the twenty items were covered during the semester. Only four were not emphasized. Therefore the remaining sixteen items, according to the instructor, were emphasized more than twice during the semester. The students confirmed that the objectives shown them were those which were included on the course syllabus and were covered during the semester. However, at least three of the seven students agreed that objectives 1, 2, 4, and 5 were given the most attention, which concurs with the instructor's assessment. The students were also confident that they had achieved the learning objectives.

In reference to the subtest items, some students identified items which were not covered during the semester. Yet those particular items conflicted with the same items many of the students felt were emphasized. At least three of the seven students thought fifteen of the twenty items were emphasized, fourteen of which corresponded with those identified by the instructor.

The relationship between these items and the objectives (1,2,4,5) which both the instructor and the students felt were emphasized was as follows: four items related to objective 1, five items related to objective 4, and one item related to objective 5. None of the items related to objective 2. Although no one interviewed selected objective

3 as being emphasized, both groups concurred that the items related to objective 3 were emphasized during the semester. The reason the objective may not have been recognized as being emphasized could be due to the general manner in which it is stated, as well as the possibility that the terminology used in the objective may not have been used during instruction and was therefore unfamiliar to both the students and the instructor.

CT 3204 - Textile Science

The instructor and five students were interviewed during individual sessions. The average score based on all students taking the posttest in this course was 65. The students' scores on the posttest were 80 (2), 60, 45, and 30.

The instructor agreed that the twelve objectives (Table 16) she was shown were included in the course syllabus for CT 3204. Subtest items related to the objectives are in Appendix C. She noted that the objectives were used to organize the course and to serve as a guide in developing tests. The objectives were also used to determine how much time would be spent on various topics. In the instructor's opinion, all objectives were covered and were well balanced. However, those which were emphasized included objectives one through four. According to the instructor, all items on the subtest were covered during the semester. The eight items relating to fiber properties were assigned a greater percentage of class time (4-5 weeks).

All of the students confirmed that a syllabus was distributed and included the objectives and that all of the objectives presented in the

Table 16. Course Objectives for CT 3204 - Textile Science

1. Define various textile terms and use them in proper context.
 2. Distinguish between natural and man-made fibers, and among various members of each class.
 3. Relate properties of the various fibers to their chemical structures.
 4. Evaluate various physical and chemical tests for fiber identification with consideration of inherent limitations.
 5. Distinguish between spun (staple) and filament yarns, complex and miscellaneous yarns.
 6. Use a yarn classification system to select a yarn type suitable for a textile product end use.
 7. Distinguish among nonwoven, woven, and knit fabrics.
 8. Relate the properties and principal uses of the various types of fabric constructions to consumer, industrial, and high performance applications.
 9. Recognize certain fabric finishes and coloration methods. Understand the chemical structures of finishes and dyes in relation to affinity for substrates and processing procedures.
 10. Recognize the mechanical, aesthetic, functional effects of processing fibers, yarns, fabrics with selected finishes and coloration and the effects of these treatments on appearance and performance.
 11. Analyze selected textile periodicals, according to specified guidelines, in relation to new developments in the textile industry.
 12. Identify the names of common fabrics by physical characterization.
-

table were covered during the semester. Those which at least two of the five students felt were emphasized included objective 1, 3, 5, 6, 7, and 12. Only objectives 1 and 3 concurred with those selected by the instructor. All of the students thought they had achieved the twelve learning objectives, but not to the same extent.

In reference to the items on the subtest, three of the students felt all of the items were covered, some more than others. Eleven items were identified by at least two of the students as being given the most emphasis in the class; six of those items were also identified by the instructor. Two of the five students identified one item which they felt had not been covered in the class. Of the six items which both the students and the instructor felt were emphasized, none related to the two objectives which both agreed were also emphasized. Rather, the six items related to objective 2 which only the instructor said was emphasized.

Interview With Seniors

Ten seniors were interviewed and asked about the relationship between the items on the exam and the information taught in the clothing and textiles curriculum. The average score for all upperclassmen taking the consolidated exam was 68.67. The scores for the seniors interviewed were 88, 84 (2), 78, 74, 73, 72, 68, 67, and 54.

The majority of the seniors felt that all of the items on the core knowledge exam had been covered in one or more courses in the curriculum. Four students agreed on four items which they felt had been emphasized in the curriculum. At least four students identified only

one item which they thought had not been covered in the curriculum. Many of the students commented on the great amount of overlap in information taught in the department courses. Most of the students, however, also felt confident about their success in department courses.

Summary

Overall instructors and students in each of the five courses agreed that all of the course objectives and items on the subtests were covered during the semester; some objectives and items were given more emphasis than others. Table 17 summarizes instructor and student responses in terms of the objectives both groups felt were emphasized and the number of subtest items related to the objectives. In most cases, the objectives which were emphasized had related test items which the students and the instructors felt had also been emphasized during the semester. In general, however, the subtests were not well balanced in terms of the items related to the objectives.

Seniors were confident that the information on the core knowledge exam was covered in the clothing and textiles curriculum. Therefore, content validity has been established: the items on the subtest were related to the covered in the five core courses.

Reliability

Hypothesis 9a: The internal consistency estimate for reliability for the subtests and for the core knowledge exam in its consolidated form will be greater than or equal to .70.

Table 17. Summary of Instructor and Student Responses

Course	Common Objectives Emphasized (Students & Instructors)	Objectives & Number of Related Items on Subtests
CT 1004	2,4	#2(4), #4(0)
CT 2404	1,4	#1(12), #4(6)
CT 2414	1	#1(7)
CT 2604	1,2,4,5	#1(6), #2(1) #4(8), #5(2)
CT 3204	1,3	#(0), #3(0)

Internal Consistency

Two statistical tests were used to evaluate the instrument for internal consistency. They were the KR-20 reliability estimate and split half reliability using Pearson product-moment correlation and the Spearman Brown Formula (Table 18). Reliability based on the KR-20 formula was higher for the CT 1004 posttest (.664) and the CT 3204 posttest (.701) than for the other three courses. Reliability based on the split half method was also higher for CT 1004 (.6864) and for CT 3204 (.7782). The reliability estimates are markedly different from each other for CT 2404, CT 2414, and CT 2604 which is probably a reflection of the nature of the two formulas; the split half estimate is more similar to a parallel forms estimate than to the KR-20 estimate.

The reliability estimates were also calculated for the instrument in its consolidated form (Table 18). The results show a high estimate of internal consistency using the split half method and the KR-20 estimate. This hypothesis was accepted for the consolidated exam only. In comparison to the subtests, the estimate for the full exam is a more realistic reliability estimate since test length influences reliability. Subtests for each class represent only 20% of the length of the instrument in its consolidated form. Therefore, reliability estimates for the shorter tests are low in comparison to those for the consolidated exam. Other factors, such as distractor effectiveness and item discriminatory power, have influenced the reliability coefficients and will be discussed later.

Table 18. Comparison of Reliability Estimates

Test	n	Split-half Estimate	KR-20 Estimate
CT 1004	57	.6864	.664
CT 2404	40	.0042	.375
CT 2414	30	.1159	.463
CT 2604	26	.5968	.292
CT 3204	12	.7782*	.701*
Consolidated	24	.8425*	.830*

*Estimate is $\geq .70$

Hypothesis 9b: The stability estimate for reliability for the test in its consolidated form will be greater than or equal to .70.

Stability

Pearson product-moment correlation was used to compare the two sets of scores from the test-retest group of upperclassmen (Table 19). The correlation (.9074) between the administration of the first exam and the administration of the second exam was positive and statistically significant at $p = .0001$, leading to the acceptance of the hypothesis. These results illustrate the ability of the exam to be a consistent measure of student knowledge of clothing and textiles core information over time.

Item Quality

Hypothesis 10: The item analysis will yield information related to the reliability of the subtests and the consolidated test. The item analysis will also relate to information gained from interviews with students and course instructors.

An item analysis was conducted for each course specific component of the core knowledge exam. The item analysis included the determination of item difficulty levels, item discriminatory power, and item distractor effectiveness. Consistent with recommendations from the literature, each group was divided into the upper and lower 27%; the same number of individuals was in each group. The difficulty level and discrimination index were calculated for each item. The difficulty level ranges from 0 to +1. The higher the number, the easier the item.

Table 19. Results of Test-Retest Stability Measure

Test n = 12	<u>M</u>	<u>SD</u>	r	p
Test	71.33	8.84		
			.9074*	.0001
Retest	74.67	7.53		

*Statistical value meets pre-established significance level.

Ebel (1972) suggests that difficulty levels between .25 and .75 contribute to test reliability, therefore this is the desired range.

In reference to the discrimination index, a high number illustrates the ability of an item to distinguish between high scorers and low scorers. The index can range from -1 to +1. A high positive number indicates that more high scorers answered the item correctly than low scorers. Indices below .30 are usually in need of revision according to Ebel (1972). The detailed results for each class and the consolidated exam are presented in Appendices D and E and will be discussed by course. Table 20 presents the average difficulty and discrimination indices for all subtests and for the consolidated form of the exam.

CT 1004

Results of the item analysis for the CT 1004 subtest are presented in Table 21. The item analysis showed that twelve items were in the acceptable range of difficulty (.25 to .75). The average difficulty index was .62. Three of the twelve were items the students and instructor thought were emphasized in class. All items reported by the instructor and students as emphasized were above the minimum acceptable level of difficulty (.25). Only two items on the subtest were below the lower limit, indicating that these items were too difficult. Conversely, the remaining six were above the .75 level, therefore too easy.

The discrimination index ranged from $-.067$ to $+.67$ with a mean of .33. Only nine of the items were above the .30 preferred level. Five items, however, were very close to the .30 cutoff and contributed

Table 20. Summary of Mean Difficulty and Discrimination Index For Subtests and Consolidated Exam

Course	Subtest		Consolidated Test	
	Difficulty <u>M</u>	Discrimination <u>M</u>	Difficulty <u>M</u>	Discrimination <u>M</u>
CT 1004	.62	.33	.73	.16
CT 2404	.54	.30	.71	.23
CT 2414	.70	.28	.73	.21
CT 2604	.67	.24	.66	.24
CT 3204	.62	.37	.65	.37
Across Courses	.63	.31	.70	.24

Table 21. Results of Item Analysis for CT 1004 - Clothing and People

Item No.	Emphasis		Diffic. Index	Discrim. Index
<u>CT 1004</u>				
10		SE	.70*	.067
11	IE	SE	.57*	.60*
12	IE	SE	.90	.20
13		SE	.70*	.47*
14		SE	.63*	.40*
15	IE	SE	.80	.40*
16	IE	SE	.73*	.27
17		SE	.60*	.67*
18	IE	SE	.73*	.40*
19		SE	.70*	.60*
20			.033	- .067
21			.73*	.13
22			.13	.27
23	IE		.50*	.60*
24	IE	SE	.77	.20
25		SE	.40*	.27
26			.90	.20
27			.77	.47*
28	IE		.87	.27
29			.33*	.27

* = Acceptable

IE = Instructor Recalled Emphasis

SE = Student Recalled Emphasis

slightly in discriminating between the high and low scorers. Eleven of the items were below the .30 level and therefore not discriminating between the high and low scorers. Of the eleven, one item had a negative index indicating that fewer of the high scorers answered the item correctly. These items will need to be evaluated and probably revised. This particular item also had a very low difficulty level (.033), which indicates that very few students selected the correct answer. That this item was not one which the instructor and students identified as being emphasized in class indicates that the students may not have been familiar with it.

Seven of the twenty items (45%) had an acceptable difficulty level and a correspondingly acceptable discriminatory index. These seven items were also selected by the instructor and/or the students as having been emphasized in class. The reliability (KR-20 = .66) showed some internal consistency, although not at the .70 desired level, and was influenced by the large number of items (14) which were either above or very close to the .30 cut off for the discrimination index. All distractors were used for fifteen items; however, for only four items were all distractors used by both the high and low scorers.

CT 2404

Results of the item analysis for CT 2404 subtest are presented in Table 22. The difficulty level of sixteen (80%) of the items on the CT 2404 subtest were in the acceptable range. The average difficulty index was .54. Seven of the fifteen items had been selected by the instructor and/or students as having been emphasized during the semester. One item

Table 22. Results of Item Analysis for CT 2404 - Basic Apparel Assembly Lecture

Item No.	Emphasis		Diffic. Index	Discrim. Index
<u>CT 2404</u>				
30			.82	0
31	IE		.91	.18
32	IE	SE	.96	.09
33	IE		.45*	.18
34			.45*	.36*
35		SE	.27*	.55*
36	IE		.68*	.27
37			.55*	.73*
38	IE	SE	.45*	.18
39			.27*	.18
40			.41*	.09
41			.59*	.27
42		SE	.68*	.45
43	IE	SE	.68*	.09
44			.45*	.55*
45			.41*	.27
46			.68*	.45*
47			.36*	.55*
48			.50*	.45*
49	IE		.18	.18*

* = Acceptable
 IE = Instructor Recalled Emphasis
 SE = Student Recalled Emphasis

which the instructor thought had been stressed was below the 25% difficulty level and was also below the 30% discrimination index. This result indicates that very few students answered correctly. The mean discrimination index was .30 and the index ranged from 0 to .73. Only eight items effectively discriminated between high and low scorers. These eight items also had acceptable difficulty levels. Two of the items which had a discrimination index above .30 were indicated by the students as emphasized during the semester. The items which the instructor indicated were emphasized did not have acceptable discrimination indices and difficulty levels indicated these items were too either too easy or too difficult.

All distractors were used for fifteen items. The items for which the distractors were not used should be revised to increase the discrimination index and decrease the difficulty index. Only five items had all distractors used by both the high and low scorers. The large percentage of ineffective items in terms of discriminatory ability and distractor effectiveness probably contributed to the low reliability estimate ($KR-20 = .375$).

CT 2414

The item analysis for the CT 2414 subtest revealed a mean discrimination index of .28 with only nine items performing above the acceptable .30 level (Table 23). One item yielded a negative discrimination index and two yielded an index of zero. The mean difficulty index was .70 with ten items performing within the acceptable range of difficulty, seven of which also discriminated effectively. The

Table 23. Results of Item Analysis for CT 2414 - Basic Apparel Assembly Laboratory

Item No.	Emphasis		Diffic. Index	Discrim. Index
<u>CT 2414</u>				
50	IE	SE	.94	.13
51	IE	SE	.75*	.25
52			.44*	.63*
53			.81	.38*
54			.81	.13
55	IE		.81	.38*
56	IE	SE	.69*	.13
57			.81	.13
58			.63*	.50*
59	IE	SE	1.00	0
60	IE		.50*	.50*
61			.88	.25
62			.56*	.88*
63			.88	.25
64	IE	SE	.63*	-.25
65	IE	SE	1.00	0
66			.31*	.38*
67		SE	.88	.25
68		SE	.31*	.38*
69			.44*	.38*

* = Acceptable
 IE = Instructor Recalled Emphasis
 SE = Student Recalled Emphasis

remaining items (10) were above the .75 difficulty level cutoff which suggests that most of the test items were too easy. One of the seven items which had both acceptable discrimination and difficulty indices was selected by students as having been emphasized in class. Two of the items which both the instructor and students felt had been emphasized during the semester had a difficulty level of 1.00 which indicates that it was an extremely easy item that should be revised. On the other hand one item which both groups thought had been emphasized, the high scorers did poorly on and will also need revision. The difficulty level for this item was .63, but the discrimination index was -.25.

All distractors were used for seven items: only one of which attracted students from both the high and low groups. The reliability estimate ($KR-20 = .467$) was low and was influenced by the low discrimination indices for more than half of the items.

CT 2604

The difficulty level for the CT 2604 subtest ranged from .21 to 1.00 with a mean of .67 (Table 24). Nine items were within the middle difficulty range. The mean discrimination index was .24, and only four items effectively discriminated between high and low scorers. These items also had acceptable levels of difficulty. Results of this nature suggest that more than half of the items were too easy.

Of the four which had both effective discrimination indices and acceptable levels of difficulty, two items had been selected by the instructor and students as having been emphasized in class. Three items which the instructor and the students reported were emphasized had

Table 24. Results of Item Analysis for CT 2604 - Introduction to the Fashion Industry

Item No.	Emphasis		Diffic. Index	Discrim. Index
<u>CT 2604</u>				
70		SE	.86	.29
71	IE	SE	.93	.14
72	IE	SE	.93	.14
73	IE		.71*	.29
74	IE		.21	.14
75	IE	SE	.43*	.29
76	IE	SE	1.00	0
77	IE	SE	.50*	.14
78			.21	.14
79	IE	SE	.57*	.57*
80	IE	SE	.86	.29
81	IE	SE	.64*	.71*
82			.50*	.43*
83			.36*	.43*
84	IE	SE	.29*	.29
85	IE	SE	1.00	0
86	IE	SE	.57*	.29
87	IE	SE	.93	.14
88	IE	SE	1.00	0
89	IE	SE	.93	.14

* = Acceptable

IE = Instructor Recalled Emphasis

SE = Student Recalled Emphasis

difficulty levels of 1.00 and were apparently too easy. One item which the instructor selected as being given emphasis in class had a difficulty level and discrimination index below the acceptable limit. This result implies that the students may not have been familiar with the item. All items which both the instructor and the students thought were emphasized had difficulty levels above the .25 lower limit.

All distractors were used for only five items which is a reflection of how easy many of the items were. Only one item attracted students from both the high and low groups. The lack of the subtest's internal consistency ($KR-20 = .292$) was influenced by the low percentage of items which actively discriminated, the lack of effective distractors, and the high difficulty indices of many items.

CT 3204

The item analysis for the CT 3204 subtest was based on a total group size of six with three each in the high and low groups (Table 25). The mean discrimination index was .37 with sixteen items contributing effectively in terms of discriminatory power. The four items which did not reach the .30 level did not discriminate at all (index = 0). The mean difficulty index was .62 with thirteen items performing within the acceptable range. Twelve of the thirteen items also discriminated effectively. Three items, however, yielded difficulty levels of +1 and a discrimination index of 0, which indicates that these were very easy items for both groups and will need to be revised. These three items were identified by students as having been given emphasis in class. All but one item selected by students as being given emphasis had difficulty

Table 25. Results of Item Analysis for CT 3204 - Textile Science

Item No.	Emphasis		Diffic. Index	Discrim. Index
<u>CT 3204</u>				
90		SE	1.00	0
91		SE	.17	.33*
92	IE	SE	.50*	.33*
93	IE	SE	1.00	0
94			.50*	.33*
95		SE	.67*	.67*
96		SE	.67*	0
97			.67*	.67*
98			.17	.33*
99		SE	.83	.33*
100			.83	.33*
101	IE		.50*	.33*
102	IE	SE	.33*	.67*
103	IE	SE	1.00	0
104	IE	SE	.50*	.33*
105	IE		.67*	.67*
106	IE	SE	.67*	.67*
107			.50*	.33*
108			.67*	.67*
109			.50*	.33*

* = Acceptable
 IE = Instructor Recalled Emphasis
 SE = Student Recalled Emphasis

levels above .25. All items selected by the instructor as being emphasized in class had difficulty levels above .25. This result indicates that the students were familiar with most of the items. All distractors were used for only two items by either the high or low group, which also is an indication of the students' familiarity with the information. The distractors will need to be revised for most of the items.

Because of the low number of students in the high and low groups, the total group (12) was considered. The mean discrimination index for the total group was .3085 with eleven (55%) items displaying acceptable discrimination indices. Ten of the eleven items also displayed acceptable difficulty levels. Although the mean discrimination index dropped for the total group, at least half of the test items were adequate in terms of discriminatory power and level of difficulty.

The reliability estimate (KR-20 = .70) for internal consistency was higher for this subtest than the other four and was influenced by the high percentage of items with acceptable difficulty levels and discriminatory power. However, a large number of items (18) had ineffective distractors, which indicates the distractors need to be evaluated and revised.

Consolidated Core Knowledge Exam

The core knowledge exam in its consolidated form (100 items) was administered to twenty-four upperclass students. Item analysis was conducted on the exam for a criterion group of fourteen students with seven each in the high and low groups. Results of the item analysis are

presented in Table 27. The mean discrimination index was .24 with twenty-three items exceeding the acceptable .30 discrimination level. Although forty-five items fell within the acceptable difficulty range, only eighteen of these yielded the minimum discrimination index of .30. As shown in Tables 26 and 27, which compares the results of the subtest with the consolidated exam, only nine of the eighteen items also had acceptable difficulty levels and discrimination indices on the subtests. Several items (28) on the consolidated exam reached an index of .29, which is very close to the minimum acceptable index. Seventeen items had a discrimination of zero and could be considered too easy. Two of the seventeen items also yielded an index of zero on the subtests. Six items yielded negative discrimination indices, one of which also yielded a negative index on the subtest. One item on the consolidated exam had a perfect (+1.00) discrimination index which indicates that all of the high scorers and none of the low scorers answered the item correctly.

Only nineteen items had all distractors used by either the high or low scorers. Of the nineteen, only two items had all distractors used by both groups. Two of the items which the seniors identified as being emphasized in the curriculum had high difficulty levels and discrimination indices below the minimum, which illustrates their familiarity with the information. Thirty three of the 100 items had difficulty levels within the acceptable range on both the subtest version and the consolidated version of the core knowledge exam test. However only eleven items on both versions of the exam had discrimination indices above .30.

Table 26. Summary of Comparison of Item Analyses for Subtests and Consolidated Exam

Course	Difficulty Level		Discrimination Ability	
	<u>M</u>	n(index = .25-.75)	<u>M</u>	n(index > .30)
CT 1004	.62	12 items	.33	9 items
CT 2404	.54	16 items	.30	8 items
CT 2414	.70	10 items	.28	9 items
CT 2604	.67	9 items	.24	4 items
CT 3204	.62	13 items	.37	16 items
Across Courses	.63	60 items	.31	69 items
Consolidated	.70	45 items	.24	23 items

Table 27. Detailed Comparison of Item Analyses for Subtests and Consolidated Exam

Item No.	Emphasis	Subtests		Discrim. Index	Consolidated Exam		Outcome
		Diffic. Index	Diffic. Index		Discrim. Index		
CT 1004							
10	SE	.70*	.067	.79	.14	Revise	
11	IE	.57*	.60*	.86	0	Subtest OK	
12	IE	.90	.20	1.00	0	Revise	
13	SE	.70*	.47*	.71*	.29	Good	
14	SE	.63*	.40*	.93	.14	Subtest OK	
15	IE	.80	.40*	.79	.14	Review	
16	IE	.73*	.27	.79	.14	Review	
17	SE	.60*	.67*	1.00	0	Subtest OK	
18	SE	.73*	.40*	.79	.43*	Good	
19	SE	.70*	.60*	.71*	.29	Good	
20		.033	-.067	.14	-.29	Key inaccurate	
21		.73*	.13	1.00	0	Revise	
22		.13	.27	.21	.14	Revise	
23	IE	.50*	.60*	.57*	.29	Good	
24	IE	.77	.20	.86	0	Revise	
25	SE	.40*	.27	.07	.14	Revise	
26		.90	.20	.86	.29	Revise	
27		.77	.47*	.71*	.57*	Good	
28	IE	.87	.27	1.00	0	Revise	
29		.33*	.27	.79	.43*	Review	

SE = Emphasis recalled by students
IE = Emphasis recalled by instructor
* = Within acceptable limits

Table 27. Detailed Comparison of Item Analyses for Subtests and Consolidated Exam (Continued)

Item No.	Subtests			Consolidated Exam			Outcome
	Emphasis	Diffic. Index	Discrim. Index	Emphasis	Diffic. Index	Discrim. Index	
CT 2404							
30		.82	0		.79	.14	Revise
31	IE	.91	.18		.71	.57*	Revise
32	IE	.96	.09		1.00	0	Revise
33	IE	.45*	.18		.86	.29	Revise
34		.45*	.36*		.50*	.43*	Very Good
35	SE	.27*	.55*	SE	.21	.43*	Subtest OK
36	IE	.68*	.27		.79	-.14	Review
37		.55*	.73*		.64*	.71*	Very Good
38	IE	.45*	.18		.07	-.14	Revise
39		.27*	.18		.57*	.29	Review
40		.41*	.09		.71*	.29	Review
41		.59*	.27		.79	.14	Review
42	SE	.68*	.45*		.93	.14	Subtest OK
43	SE	.68*	.09		.50*	.43*	Review
44	IE	.45*	.55*		.93	-.14	Review
45		.41*	.27		.79	.14	Review
46		.68*	.45*		.93	.14	Subtest OK
47		.36*	.55*		1.00	0	Review
48		.50*	.45*		.93	.14	Subtest OK
49	IE	.18	.18		.50*	.71*	Full test OK

SE = Emphasis recalled by students
IE = Emphasis recalled by instructor
* = Within acceptable limits

Table 27. Detailed Comparison of Item Analyses for Subtests and Consolidated Exam (Continued)

Item No.	Subtests			Consolidated Exam			Outcome	
	Emphasis	Diffic. Index	Discrim. Index	Emphasis	Diffic. Index	Discrim. Index		
CT 2414								
50	IE	SE	.94	.13	SE	.86	.29	Revise
51	IE	SE	.75*	.25		.50*	.43*	Good
52			.44*	.63*		.50*	-.14	Review
53			.81	.38*		.86	.29	Revise
54			.81	.13		.86	0	Revise
55	IE		.81	.38*		.64*	.43*	Good
56	IE	SE	.69*	.13		.71*	.29	Review
57			.81	.13		.93	.14	Revise
58			.63*	.50*		.86	.29	Subtest OK
59	IE	SE	1.00	0		.93	.14	Revise
60	IE		.50*	.50*		.86	.29	Subtest OK
61			.88	.25		.93	.14	Revise
62			.56*	.88*		.57*	.29	Good
63			.88	.25		.93	.14	Revise
64	IE	SE	.63*	-.25		.43*	0	Revise
65	IE	SE	1.00	0		1.00	0	Revise
66			.31*	.38*		.64*	.14	Subtest OK
67		SE	.88	.25		.86	.29	Revise
68		SE	.31*	.38*		.29*	.29	Subtest OK
69			.44*	.38*		.50*	.43*	Very Good

SE = Emphasis recalled by students
 IE = Emphasis recalled by instructor
 * = Within acceptable limits

Table 27. Detailed Comparison of Item Analyses for Subtests and Consolidated Exam (Continued)

Item No.	Subtests		Consolidated Exam			Outcome
	Emphasis	Diffic. Index	Discrim. Index	Emphasis	Diffic. Index	
CT 2604						
70		SE	.86	.29	.79	Revise
71	IE	SE	.93	.14	.64*	Full test OK
72	IE	SE	.93	.14	1.00	Revise
73	IE		.71*	.29	.93	Revise
74	IE		.21	.14	.36*	Full test OK
75	IE	SE	.43*	.29	.21	Revise
76	IE	SE	1.00	0	1.00	Revise
77	IE	SE	.50*	.14	.50*	Good
78			.21	.14	.29*	Revise
79	IE	SE	.57*	.57*	.71*	Good
80	IE	SE	.86	.29	.50*	Revise
81	IE	SE	.64*	.71*	.71*	Good
82			.50*	.43*	.29*	Subtest OK
83			.36*	.43*	.50*	Good
84	IE	SE	.29*	.29	.64*	Key inaccurate
85	IE	SE	1.00	0	.86	Revise
86	IE	SE	.57*	.29	.43*	Good
87	IE	SE	.93	.14	1.00	Revise
88	IE	SE	1.00	0	1.00	Revise
89	IE	SE	.93	.14	.93	Key inaccurate

SE = Emphasis recalled by students
IE = Emphasis recalled by instructor
* = Within acceptable limits

Table 27. Detailed Comparison of Item Analyses for Subtests and Consolidated Exam (Continued)

Item No.	Subtests		Consolidated Exam			Outcome
	Emphasis	Diffic. Index	Discrim. Index	Emphasis	Diffic. Index	
CT 3204						
90	SE	1.00	0	.93	.14	Revise
91	SE	.17	.33*	.14	.29	Revise
92	IE	.50*	.33*	.43*	0	Subtest OK
93	IE	1.00	0	.93	.14	Revise
94		.50*	.33*	.64*	.43*	Very Good
95	SE	.67*	.67*	.64*	.29	Good
96	SE	.67*	0	.71*	.57*	Review
97		.67*	.67*	.79	.14	Subtest OK
98		.17	.33*	.57*	.29	Review
99	SE	.83	.33*	.71*	.29	Review
100		.83	.33*	.71*	.29	Review
101	IE	.50*	.33*	.86	.29	Subtest OK
102	IE	.33*	.67*	.71*	.57*	Very Good
103	SE	1.00	0	.57*	.29	Revise
104	SE	.50*	.33*	.57*	.29	Good
105	IE	.67*	.67*	.64*	.71*	Very Good
106	IE	.67*	.67*	.71*	.29	Good
107		.50*	.33*	.64*	.43*	Very Good
108		.67*	.67*	.64*	.57*	Very Good
109		.50*	.33*	.50*	1.00*	Very Good

SE = Emphasis recalled by students
 IE = Emphasis recalled by instructor
 * = Within acceptable limits

Results of the item analyses for the subtest and the consolidated exam has shown the relationship between the student and faculty interviews and the item analysis. For example, many of the items which the groups recalled as being given the most attention in class had difficulty levels above the minimum acceptable limit and low discrimination indices. The item analyses were also related to the reliability estimates of the subtests, leading to the acceptance of this hypothesis. The outcome of the item analysis is presented in the last column in Table 27 and shows that a number of items are in need of revision. Items for which all distractors were not used will need to be revised to include more effective distractors which will reduce the item difficulty indices.

Characteristics of Student Involvement

Hypothesis 11: Student outcomes will be negatively affected by involvement in student organizations.

Six sub-hypotheses were formulated relative to the relationship between the level of involvement in student organizations and the various outcome measures for students in the subtest sample who took the posttest and upperclass clothing and textiles majors who took the consolidated form of the exam. A series of independent sample t-tests were conducted to test the relationship.

Hypothesis 11a: Students in the subtest sample who are involved in fewer than three organizations will score higher on the posttest version of the course-specific subtest.

A t-test analysis was conducted across courses to test the relationship between the level of student involvement in organizations and the performance on the posttest (Table 28). Comparison of the posttest means for the group participating in fewer than three organizations and three or more organizations was not in the direction anticipated and was not significant, thus leading to a rejection of this hypothesis. Although students who were involved in three or more organizations scored higher on the posttest than those participating in fewer than three organizations, the difference was not significant.

Hypothesis 11b: Students in the subtest sample who are involved in fewer than three organizations will score higher on the comprehensive final examination in the core courses.

A t-test analysis was conducted on all subtests combined to test the relationship between the level of student involvement in organizations and the grades on the final examinations in four of the core courses (Table 28). The difference between mean examination scores for the two groups (those participating in fewer than three organizations and those participating in three or more organizations) was not in the direction anticipated and was not significant at the .05 level. Therefore, this hypothesis was rejected. Students who were involved in three or more organizations had a slightly higher mean course grade; a larger sample size may have revealed a significant difference.

Hypothesis 11c: Students in the subtest sample who are involved in fewer than three organizations will receive higher grades in the core courses.

Table 28. Relationship Between Outcomes Scores and Number of Organizations for the Subtest Sample

Number of Organizations	Posttest			Course Grade			Exam		
	n	M	t	n	M	t	n	M	t
3 or more	34	65.88		34	2.78		19	76.38	
			-1.20			-.15			-.19
Fewer than 3	128	63.13		128	2.77		89	75.94	

A t-test analysis was conducted on all subtests combined to test the relationship between level of student involvement in organizations and performance in the core courses based on final grades received (Table 28). Results of the t-test indicated that students involved in fewer than three organizations had a lower mean core course grade. Because this was not the anticipated outcome and the difference between the means was not significant at the .05 level this hypothesis was rejected.

Hypothesis 11d: Upperclass clothing and textiles majors who are involved in fewer than three organizations will receive a higher score on the consolidated form of the test.

A t-test analysis was conducted to test the relationship between the level of student involvement in organizations and the performance on the core knowledge exam (Table 29). The means for both groups were the same; therefore, the hypothesis was rejected.

Hypothesis 11e: Upperclass clothing and textiles majors who are involved in fewer than three organizations will have higher overall QCAs.

A t-test analysis was conducted to test the relationship between the level of student involvement in organizations and the performance in school as measured by the quality credit average (Table 29). Contrary to expectations, upperclass clothing and textiles majors who were involved in fewer than three organizations had lower mean overall QCAs; a difference which was not significant. Therefore this hypothesis was rejected.

Hypothesis 11f: Upperclass clothing and textiles majors who are involved in fewer than three organizations will have higher QCAs in the major.

Table 29. Relationship Between Outcomes Scores and Numbers of Organization for Upperclass Majors

Number of Organizations	<u>Test Score</u>		<u>Overall QCA</u>		<u>Major QCA</u>	
	<u>M</u>	<u>t</u>	<u>M</u>	<u>t</u>	<u>M</u>	<u>t</u>
3 or more (n = 12)	68.67		2.68		3.17	
		0		-.96		-1.69
Fewer than 3 (n=12)	68.67		2.52		3.03	

A t-test analysis was conducted to test the relationship between the level of student involvement in organizations and the performance in the major as measured by the quality credit average for courses in the major (Table 29). Contrary to the hypothesis upperclass clothing and textile majors who were involved in fewer than three organizations had lower QCAs in the major; a difference which was not significant. Therefore this hypothesis was rejected.

Although the t-test analyses for the six hypotheses did not provide evidence of a significant difference on outcomes measures based on the level of student involvement, an examination of the means in Tables 28 and 29 reveal that students involved in three or more organizations had higher means for all outcome measures, except the consolidated exam, than those involved in fewer than three organizations. This implies that the more students are involved in extra-curricular activities (organizations) the greater the academic benefits.

Hypothesis 12: Student outcomes will be negatively affected by involvement in paid employment.

Six sub-hypotheses were formulated relative to the relationship between the level of employment and various measures of student achievement. Independent sample t-tests were conducted to test the relationship between employment and student outcomes.

Hypothesis 12a: Students in the subtest sample who work more than ten hours per week will score lower on the posttest.

A t-test analysis was conducted to test the relationship between the level of student employment and the performance on the posttest (Table 30). In contrast to the hypothesized relationship, students in

Table 30. Relationship Between Outcomes Scores and Hours of Employment for Subtest Sample

Employed more than 10 hours	Posttest			Course Grade			Exam		
	n	<u>M</u>	<u>t</u>	n	<u>M</u>	<u>t</u>	n	<u>M</u>	<u>t</u>
Yes	52	64.13		52	2.77		37	77.07	
			.27			-.2.47E-02			.66
No	102	63.48		102	2.77		69	75.66	

the subtest sample who worked more than ten hours per week had a higher mean on the posttest; a difference which was not significant at the .05 level. Therefore this hypothesis was rejected.

Hypothesis 12b: Students in the subtest sample who worked more than ten hours per week will receive lower final grades in the core courses.

A t-test analysis was conducted to test the relationship between the performance in the core courses as measured by final course grades and the level of employment (Table 30). The difference between the means for course grades of those working more than ten hours per week and those not working more than ten hours per week was not significant at the .05 level. Therefore this hypothesis was rejected.

Hypothesis 12c: Students in the posttest group of the subtest sample who worked more than ten hours per week will earn lower examination grades in the core courses.

A t-test analysis was conducted to test the relationship between the level of employment and the performance on the final examinations in the core courses (Table 30). The t-test results revealed that students in the posttest group who worked more than ten hours per week had a higher mean for the final examination grade; a difference which was not significant at the .05 level. Therefore, this hypothesis was rejected.

Hypothesis 12d: Upperclass clothing and textiles majors who work more than ten hours per week will have lower scores on the consolidated form of the test.

A t-test analysis was conducted to test the relationship between the level of employment and the performance on the core knowledge exam (Table 31). Although upperclass clothing and textiles majors who worked more than ten hours per week had a lower mean score on the consolidated

Table 31. Relationship Between Outcomes Scores and Hours of Employment for Upperclass Majors

Employed more than 10 hours	<u>Test</u>		<u>Major QCA</u>		<u>Overall QCA</u>	
	<u>M</u>	<u>t</u>	<u>M</u>	<u>t</u>	<u>M</u>	<u>t</u>
Yes (n = 16)	66.44	-1.65	3.0	-1.69	2.5	-2.13*
No (n = 8)	73.12		3.3		2.8	

*p < .05

exam than those who did not work more than ten hours per week, the difference was not significant. Therefore this hypothesis was rejected.

Hypothesis 12e: Upperclass clothing and textiles majors who work more than ten hours per week will have lower QCAs in the major.

A t-test analysis was conducted to test the relationship between the level of employment and the performance in the major as measured by quality credit average for courses in the major (Table 31). The t-test revealed a lower mean QCA in the major for upperclass clothing and textiles majors who worked more than ten hours per week; a difference which was not significant. Therefore this hypothesis was rejected.

Hypothesis 12f: Upperclass clothing and textiles majors who work more than ten hours per week will have lower overall QCAs.

A t-test analysis was conducted to test the relationship between the level of employment and the performance in school as measured by overall QCA (Table 31). As expected, upperclass clothing and textile majors who worked more than ten hours per week had a lower mean overall QCA than those who did not work more than ten hours per week; the difference between the two means was significant at the .05 level ($t = -2.13$). Therefore, this hypothesis was accepted.

For the subtest sample, students had higher means on the posttest and the comprehensive final exam than those who did not work more than ten hours per week. Although the differences were not significant, this implies that employment may have a positive rather than a negative effect on student academic performance. On the other hand, for the upperclass subjects, students who worked more than ten hours per week

had lower means on all outcomes measures than those who did not work more than ten hours per week. T-test results were significant for one measure, which suggests that as students progress through the curriculum, employment may have a negative effect on academic performance. Certainly, a larger sample size is needed to confirm or negate these implications.

Hypothesis 13: Student outcomes will be positively affected by the number of hours involved in school related activities.

Six sub-hypotheses were formulated. A series of one-way analyses of variance were conducted to test the relationship between various outcomes measures and involvement in school-related activities. Level of involvement in school-related activities was grouped in four categories according to hours spent: fewer than 20 hours, 21-29 hours, 30-39 hours, and 40 hours or more.

Hypothesis 13a: Students from the subtest sample who spend more hour per week on school-related activities will score higher on the posttest.

A one-way analysis of variance was conducted to test the relationship between hours spent on school-related activities and posttest scores for those taking the subtests (Table 32). Although the posttest means for the groups spending the least amount of hours on school-related activities were lower than the group spending the greatest number of hours, the differences were not significant at the .05 level. Therefore this hypothesis was rejected.

Hypothesis 13b: Students from the subtest sample who spend more hours per week on school-related activities will receive higher grades in the core courses.

Table 32. Relationship Between Outcomes Scores and Hours Spent on School-Related Activities for the Subtest Sample

Measure	Hours Spent on Activities			
	< 20 hours	21-29 hours	30-39 hours	40 or more hours
<u>Posttest</u>				
N	60	52	26	13
Mean	63.09	63.65	63.65	66.92
Std. Error	1.78	1.96	2.77	3.92
	F=.26			
<u>Course Grade</u>				
N	60	52	26	13
Mean	2.65	2.81	2.8	3.18
Std. Error	9.92E-02	.109	.154	.218
	F=1.73			
<u>Exam Grade</u>				
N	40	38	17	11
Mean	75.19	75.83	76.97	79.45
Std. Error	1.75	1.79	2.68	3.33
	F=.47			

A one-way analysis of variance was conducted to test the relationship between hours spent on school related activities and grades received in the core courses for those taking the subtests (Table 32). Results of the analysis revealed that mean course grades increased as number of hours spent on school-related activities increased, however the difference between the means failed to reach significance ($F=1.73$, $df=3/150$). Therefore this hypothesis was rejected.

Hypothesis 13c: Students from the subtest sample who spend more hours per week on school-related activities will earn higher grades on the final examinations in the core courses.

A one-way analysis of variance was conducted to test the relationship between hours spent on school-related activities and grades earned on the final examinations (Table 32). Group means on the final examination increased as number of hours spent on school-related activities increased, however, the difference between the means failed to reach significance at the .05 level ($F=.47$, $df=3/102$). Therefore this hypothesis was rejected.

Hypothesis 13d: Upperclass clothing and textiles majors who spent more hours per week on school-related activities will receive a higher score on the core knowledge exam.

A one-way analysis of variance was conducted to test the relationship between hours spent on school-related activities and core knowledge exam score (Table 33). The analysis revealed a higher mean for the group that spent 30 to 39 hours per week on school-related activities than those spending fewer hours; however, the difference between the means was not significant at the .05 level ($F=2.43$, $df=3/20$). Therefore this hypothesis was rejected. The difference

Table 33. Relationship Between Outcomes Scores and Hours Spent on School-Related Activities for Upperclass Majors

Measure hours	Hours Spent on Activities			
	< 20 hours	21-29 hours	30-39 hours	40 or more
N	9	9	5	1
<u>Test Score</u>				
Mean	66.89	67.83	77.99	54.99
Std. Error	2.97	2.97	3.98	8.91
	F=2.43			
<u>Overall QCA</u>				
Mean	2.60	2.52	2.85	2.1
Std. Error	.127	.127	.171	.382
	F=1.44			
<u>Major QCA</u>				
Mean	3.05	3.03	3.38	2.77
Std. Error	.135	.135	.181	.404
	F=1.18			

between the means was significant at the .10 level ($p=.0952$). The group that spent 40 hours or more on school-related activities had only one subject in this cell, therefore the mean was lower than the other groups. A larger number of subjects in this cell, as well as the other cells may have revealed a significant difference.

Hypothesis 13e: Upperclass clothing and textiles majors who spend more hours per week on school-related activities will have a higher overall QCA.

A one-way analysis of variance was conducted to test the relationship between hours spent on school-related activities and school performance as measured by the overall QCA (Table 33). The mean for overall QCA was highest for the group that spent 30 to 39 hours per week on school-related activities; however, results of the one-way ANOVA were not significant at the .05 level ($F=1.44, df=3/20$). Therefore, this hypothesis was rejected. A larger sample size may yield significant results.

Hypothesis 13f: Upperclass clothing and textiles majors who spend more hours per week on school-related activities will have a higher QCA in the major.

A one-way analysis of variance was conducted to test the relationship between hours spent on school-related activities and performance in the major as measured by QCA in major courses (Table 33). Although the mean for QCA in the major was highest for the group that spent 30 to 39 hours per week on school-related activities, results of the analysis of variance were not significant at the .05 level ($F=1.18, df=3/20$). Therefore, this hypothesis was rejected. A larger sample size may yield significant results.

For the subtest subjects, as the number of hours spent on school-related activities increased, the means on all outcomes measures also increased slightly. The mean was highest for those who reported spending forty hours or more per week on school-related activities. For the upperclass subjects, the mean was highest on all outcomes measures for those who spent thirty to thirty-nine hours per week on school-related activities. There was not a consistent increase in the group means for the upperclass subjects; this could be attributed to the unequal cell sizes. The differences between the means were not significant which could be a reflection of the small sample size. However, these results imply that the more hours spent on school-related activities the better the student's academic performance.

Chapter Summary

The sample for this study was comprised of students enrolled in five core courses in the Department of Clothing and Textiles who were administered the course specific component of the core knowledge exam, upperclass clothing and textiles majors who had taken at least four of the five core courses, underclass clothing and textiles majors who had taken two or fewer core courses, and faculty who taught the courses. The upperclass majors and the underclass majors were administered the exam in its consolidated (100 items) form.

When administered as a pretest on the first day of classes, the core knowledge exam was not a predictor of course grades. Therefore predictive validity was not demonstrated. However, when administered as

a posttest on the last day of classes, the exam was able to predict course grades for CT 3204 and across all courses, thus demonstrating concurrent validity.

Content validity was demonstrated by comparing pretest scores with posttest scores; the comparison showed a significant increase in score from pretest to posttest. Interviews with a subset of students who took the subtests and the consolidated exam and with faculty also confirmed the exam's content validity.

Construct validity was demonstrated for the subtests combined and for CT 2404 by significant correlations between posttest scores and comprehensive final examination grades in four of the five core courses. The exam's capacity to measure value-added knowledge was demonstrated by significant differences between the means for the pretest and the posttest for all courses and for underclassmen and upperclassmen on the consolidated form of the exam. These results also provided further evidence that the construct, knowledge of clothing and textiles core information, was measured by the exam. For upperclassmen, quality credit average was associated with exam scores, which was also a measure of the construct.

The reliability of the exam was evaluated using two internal consistency estimates and an estimate of stability. Both types of reliability estimates (internal consistency and stability) for the consolidated exam reached or surpassed .70. It was determined through item analysis that the reliability estimate of internal consistency was

related to item discriminatory ability, difficulty level, and distractor effectiveness. Items in need of revision were identified.

The relationships between the level of involvement in employment, in student organizations, and in school related activities with the various measures of student achievement were also considered in this study. Although most of the results failed to reach the level of significance, a relationship was apparent. Students in the subtest sample who worked more than ten hours per week had higher means on all outcomes measures than those who did not work more than ten hours per week. However, the opposite was the case for the upperclass students. The hypothesis related to the level of involvement in employment and overall QCA for upperclassmen was accepted. Results showed that upperclass clothing and textiles majors who worked more than ten hours per week had lower overall QCAs than those who did not work more than ten hours per week. This result indicates that as the curriculum increases in difficulty, employment may have a negative effect on student academic performance.

Positive effects on student academic performance seemed to result from increased involvement in student organizations and in school-related activities. Although the differences between the means were not significant, students who were involved in more than three organizations and students who spent 30 to 39 hours on school related activities had higher means on all but one of the outcomes measures. A larger sample size may be needed to reveal more meaningful results.

CHAPTER V

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

This chapter will present a summary of the procedures and results of the study. Conclusions based on the results will be presented and recommendations for improving the instrument, for using the instrument, and for further research will also be presented.

Summary

The purpose of this study was to determine the validity and reliability of a program-specific assessment instrument designed to assess student mastery of core knowledge in the discipline. Additionally, the purpose was to determine the effectiveness of the instrument for ascertaining student outcomes from the value-added perspective and to determine whether performance on various outcomes assessment was affected by student characteristics. The instrument, referred to as the Clothing and Textiles Core Knowledge Exam, was evaluated for content validity, criterion-related validity, construct validity, reliability, and item quality. This study evolved from the recent emphasis on measuring the quality of the undergraduate experience. Since the majority of student time is spent in one program area, the development and use of assessment instruments related to instructional objectives and course content within the major is an essential component of student outcomes assessment. The evaluation of

the instrument developed by faculty in the Department of Clothing and Textiles to fulfill university assessment directives was the major focus of this study.

The instrument consists of 100 multiple choice items relating to the five core courses. For portions of this study, the instrument was divided into five subtests (course-specific components) each containing twenty items. The subtests were administered to the students enrolled in each of the five core courses as a pretest and a posttest. The consolidated form of the exam was administered to upperclass and underclass clothing and textiles majors. Interviews were conducted with subsets of students and with faculty who taught the courses.

Thirteen major hypotheses were formulated to evaluate the quality of the instrument. The hypotheses were based on the objectives of the study. Data analysis procedures included descriptive statistics, correlational statistics, and statistical tests of differences.

Pretest Influence

The first hypothesis was formulated to establish that the posttest scores of the groups given the pretest were not biased as a result of prior exposure to the subtest. Results of the t-test between posttest scores of those taking the subtests as a pretest and as a posttest only were not significant, indicating that the posttest scores of the pretest groups were not influenced by the pretest.

Criterion-related Validity

The purpose of the second hypothesis was to determine whether the exam when administered as a pretest could predict course grades.

Results of the analysis indicated that the exam was not an effective predictor of course grades when given as a pretest. However, based on the findings of the third hypothesis, the exam was shown to be an effective predictor of course grades when given as a posttest. Although predictive validity was not established, concurrent validity was established based on the findings of the third hypothesis.

Construct Validity

The fourth hypothesis examined the relationship between student scores on the comprehensive final exams in each course and posttest scores from the course specific components of the core knowledge exam. This hypothesis was based on Payne's (1974) suggestion for establishing construct validity which involved correlating the outcomes from two instruments which measure the same construct; the construct in this research was knowledge of core information. The results of individual course analyses revealed a positive and significant correlation for only one course; however, the results of the analysis across courses was also positive and significant. These results provided evidence that the core knowledge exam does possess construct validity.

The value-added capability and the construct validity of the exam was established in the fifth hypothesis by comparing the pretest scores and posttest scores of the groups taking both the pretest and the posttest. Although the gains in knowledge as a result of the courses varied, each course contributed to student academic development in clothing and textiles.

Since quality credit average (QCA) is often used as an indicator of knowledge acquired, scores on the core knowledge exam were correlated with QCA in the major courses for seniors. For the sixth hypothesis, there was a positive and significant correlation between QCA and test scores. Students with high QCA's received high scores on the consolidated exam.

The seventh hypothesis tested the assumption that students who have completed most of the curriculum had more knowledge of core content information than those who have had no courses or had completed only a few courses. Comparison of scores from the core knowledge exam for upperclass majors and underclass majors confirmed the assumption and provided further evidence of construct validity.

Content Validity

Content validity, which was the focus of the eighth hypothesis, was established by comparing subtest items to course objectives and interviewing the instructors of the core courses and a subset of students from each of the core courses. The purpose was to determine if the information on the subtest was taught in class. Seniors who had taken the consolidated form of the exam were also interviewed regarding the correlation between items on the exam and information taught in the courses. Interviews with instructors and students from each of the courses revealed that all of the course objectives and items on the subtests were taught. Both groups identified subtest items and objectives which were emphasized. The seniors also felt most of the information on the exam had been covered in the curriculum, if not in

the core courses. Exam items as they related to each of the objectives were placed in a table in Appendix C. Content validity was also established based on the significant increase in subtest scores from pretest to posttest (Hypothesis 5).

Reliability

Although validity is an essential component of an instrument, reliability, as demonstrated in the ninth hypothesis, is just as important. Therefore, the internal consistency of the core knowledge exam was estimated using the split-half procedure and the Kuder-Richardson 20 formula. The author, following the recommendations of Burns (1985) set the reliability cutoff at .70. Resulting coefficients for the two procedures were above the cutoff for CT 3204 and for the consolidated form of the exam. The stability of the instrument was also estimated using test-retest. The correlation between scores from two administrations of the exam was above .70 and highly significant, which suggests that the exam is reliable over time in terms of measuring core knowledge.

Item Quality

To determine what contributed to the reliability estimates for each of the subtests and for the exam in its consolidated form, an item analysis was conducted. The item analysis was also conducted to determine whether the quality of the exam, represented by the discrimination index and the difficulty index, would be related to information gained from the interviews. Analysis of the difficulty and discrimination indices and the effectiveness of the item distractors

revealed a pattern which related to the reliability estimate. Comparison of the difficulty and discrimination indices with items which students and the instructors recalled as being emphasized revealed a relationship between high difficulty indices, low discrimination indices and items emphasized in class. The resulting indices on the subtests were similar to those on the consolidated form of the test taken by the upperclassmen.

Characteristics of Student Involvement

Astin (1985b) and Pace (1984b) both espouse the idea that student involvement and/or student experiences in college affect academic outcomes. To test this idea, three major hypotheses were formulated to determine the relationships between various measures of student outcomes (posttest scores on the core knowledge exam, scores on the consolidated exam, final exam grades in core courses, final grades in core courses, overall quality credit average, and QCA in major) and student involvement in paid employment, student organizations, and school related activities. As a result of the small sample sizes, the findings for most of the hypotheses were not significant.

The one significant finding revealed that upperclass majors who worked more than ten hours per week had lower overall QCAs than those who did not work more than ten hours per week. The other hypotheses, although not significant, revealed a pattern among students. For instance, students from the subtest sample who spent more hours on school related activities received higher scores on the posttest and on final examinations in the courses. Involvement in more than two student

organizations appeared to positively affect student academic performance, which supports Astin's (1985b) contention that the more students are involved in the college experience, the greater their development.

Conclusions

The major objective of this study was to evaluate a program-specific, faculty developed assessment instrument for various forms of validity, reliability, and item quality. The instrument, referred to as the Clothing and Textiles Core Knowledge Exam, was also evaluated for its ability to determine value-added knowledge.

The results of the study demonstrated that the content of the exam was valid and related to information emphasized in the core courses. The item analyses, however, revealed items which were weak in their ability to discriminate between high scorers and low scorers and which were extremely easy. These findings suggest that the instrument needs to be revised. The results of the study also demonstrated the exam's ability to measure the construct, knowledge of clothing and textiles core information. Since course grades should represent the student's level of achievement of course objectives and final examinations, which are comprehensive, and should represent information covered in the course, the significant correlation of the core knowledge exam with course grades and comprehensive exam grades supported the premise that the core exam measures knowledge related to these core courses.

Astin (1985b) suggests that the quality of a program should be measured by its ability to add value to or increase the student's entering level of knowledge. The increase can only be determined by measuring and comparing the student's entering level of knowledge with the student's knowledge level at the end of the program. In this study, the significant increase between pretest and posttest scores on the course-specific components of the exam demonstrated the ability of the exam to measure value-added knowledge. The magnitude of the increase for each course provides useful information regarding course outcomes. Instructors may want to consider what can be done to ensure the desired level of increase within a course. Comparison of scores on the consolidated form of the exam for upperclass majors who had completed most of the curriculum and underclass majors beginning the curriculum provided evidence that students completing the curriculum have more subject-matter knowledge than entering students.

The exam was not an effective predictor of course outcomes when given as a pretest; however, the initial goal was not to use the exam as a pretest to predict outcomes, but as a posttest to measure knowledge at the end of the curriculum. On the other hand, the results of this study indicate that the exam may serve a more useful purpose if used as a value-added measure rather than an outcomes-only measure. Valuable information could be gained for the student and the department if the exam is given to freshmen upon entering the program and upon completion of the program.

Although the internal consistency of the instrument in its consolidated form and for one subtest was above .70, the lack of reliability for four of the subtests and results of the item analysis indicates the need to revise the items to improve reliability of the parts and the whole exam. For some courses there were objectives which were not represented by items. Therefore more balance is needed to improve the instrument's reliability. The ability of the instrument to be consistent over repeated trials was demonstrated by a stability estimate of .90.

Knowledge of outcomes alone is not as useful as knowing what factors affect the level or magnitude of the outcomes. Pace (1984b) and Astin (1985b) identified various factors that affect student outcomes. Those which were included in this study were employment, involvement in student organizations, and involvement in school related activities, such as studying and library work. The results of the hypothesis tests were significant for only one hypothesis. However, a comparison of the results of various outcomes provided some evidence that employment hours, number of organizations involved in, and number of hours spent on school related activities are related to academic performance. Contrary to expectations, students who were involved in three or more organizations performed better on the posttest, achieved higher course grades and examination grades, had an overall QCA that was higher, and a higher QCA in the major than those involved in fewer than three organizations. These findings appear to support Astin's (1985b) premise that the more students are involved in their educational experience, the

greater the outcomes. Involvement in paid employment affected achievement for the subtest sample differently than for upperclass majors. Mean scores on the outcomes measures were higher for those in the subtest sample who worked more than ten hours per week. However, for the upperclass majors the opposite pattern occurred. This disparity may indicate that during the early stages of the college experience, it is easier to balance employment and academics. But as the student progresses through a program and courses get increasingly difficult, employment has a negative effect on academics.

In conclusion, the instrument in its present form is a valid and reliable measure of student knowledge of core information in clothing and textiles. Nonetheless, the quality of the items must be improved if the instrument is to be used in the future. Use of the instrument provides valuable information for program improvement and for a better understanding of the student's progression in the program and factors which may affect that progression.

Recommendations

Based on the results of this study, recommendations will be made for improving the instrument, for using the instrument, and for further research.

Recommendations for Improving the Instrument

As a result of interviews with students and the item analyses for the subtests and for the consolidated form of the exam, the author recommends that the core course objectives be reviewed and revised so

they are measurable and succinct. It is also recommended that more items are added to the exam to represent all objectives in a balanced form and some items are deleted for objectives which are overrepresented. The items which were too easy and were weak in their ability to discriminate should be revised to improve the instrument's reliability and overall quality.

Recommendations for Using the Revised Instrument

Once the instrument is revised, it may be used to conduct a longitudinal study of students in the department. For instance the exam could be administered to students entering the department in their freshman year and again to the same students as seniors to determine the value-added as a result of the clothing and textiles curriculum. The instrument could also be administered to alumni to determine how long the knowledge acquired in the department is retained, how it is used in their careers, what information in particular is used, and for what kinds of jobs. Another recommendation for using the instrument is to administer it to graduate students to compare their level of core knowledge with those of undergraduates. Results from the present study and subsequent studies using the core knowledge exam should be used to make program improvements.

Recommendations for Further Research

Based on the results of this study, including suggested revisions for the instrument, the following recommendations for future research are made.

1. A study could be conducted to determine the relationship between student affective outcomes and student academic outcomes.
2. A study could be conducted to determine the relationship between various forms of student involvement with student academic achievement using a larger sample.
3. A study could be conducted to determine the relationship between achievement in the option areas and achievement in the core curriculum.
4. A study could be conducted to compare the benefits of value-added information versus outcomes only information.
5. A study could be conducted to investigate the feasibility of a national core curriculum.
6. A study could be conducted to investigate the relationship between entering student characteristics (SAT scores) and academic outcomes (QCA, scores on the core knowledge exam, and course grades).
7. The model used in this study could be used by other departments to develop a similar program-specific measure.

Recommendation from Advisory Committee

As a result of the high pretest scores and the information reported by students in the interviews, the department should identify the information student's are already familiar with when they enter the core courses. This could be accomplished by identifying the test items which the students answered correctly on the pretest or by thoroughly examining the content of the core courses for repetitive information to determine what information is being repeated and to what degree. This

information should be used to revise the information taught in the core courses.

This research represents a program-based approach for assessing student outcomes and provides a method for developing a program-based assessment instrument for measuring specific cognitive outcomes. Moreover, this research illustrates how assessment can be used to reveal program strengths and weaknesses in an effort to improve the curriculum. The results of this study indicate the need for educators to be cognizant of factors which may affect student learning. Within this study these factors included student prior knowledge of course information and student involvement in organizations, employment, and school related activities. An awareness of these factors prior to the learning experience may determine the need to adjust the instructional approach so that more value will be added to the students' knowledge base.

BIBLIOGRAPHY

- Astin, A.W. (1985a). Involvement: The cornerstone of excellence. Change, 17(4), 35-39.
- Astin, A. W. (1985b). Achieving educational excellence. San Francisco: Jossey-Bass Publishers.
- Astin, A. W. (1987). Assessment, value-added, and educational excellence. New Directions for Higher Education, No. 59, 89-107.
- Balderston, R.E. (May 1970). Thinking about the outputs of higher education. Ford Foundation Program for Research in University Administration.
- Banta, T.W. (1985). Use of outcomes information at the University of Tennessee, Knoxville. Assessing Educational Outcomes, No. 47, 19-32.
- Banta, T. W. & Moffett, M. S. (1987). Performance funding in Tennessee: Stimulus for program improvement. New Directions for Higher Education, No. 59, 35-43.
- Belcher, M. J. (1987). Value-added assessment: College education and student growth. New Directions for Community Colleges, No. 59, 31-38.
- Bloland, R. M., & Michael, W. B. (1984). A comparison of the relative validity of a measure of Piagetian cognitive development and a set of conventional prognostic measures in the prediction of the future success of ninth and tenth grade students in algebra. Educational and Psychological Measurement, 44(4), 925-43.

- Bok, D. (1986). Toward higher learning: The importance of assessing outcomes. Change, 18(6), 18-27.
- Boyer, E. L. (1987). College: The undergraduate experience in America. New York: Harper & Row, Publishers.
- Boyer, C. M., Ewell, P. T., Finney, J. E., & Mingle, J. R. (1987, March). Assessment and outcomes measurement: A view from the states. AAHE Bulletin, 8-12.
- Burns, R. L. (1985). Guidelines for developing and using licensure tests. In J. C. Fortune (Ed.), Understanding testing in occupational licensing (pp. 15-44). San Francisco: Jossey-Bass, Inc. Publishers.
- Calhoun, J. G., Wolf, F. M., Fredric, M., Maxim, B. R., & Davis, W. K. (1985). Validity coefficients of clinical competence on NBME Part III examination. Journal of Medical Education, 60(3), 189-192.
- Campbell, D. T. & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Chicago: Rand McNally College Publishing Company.
- Ciereszko, A. A. (1987). Mandated testing in Florida: A faculty perspective. New Directions for Higher Education, No. 59, 67-73.
- Conrad, C. F. & Wilson, R. F. (1985). Academic program reviews: Institutional approaches, expectations, and controversies. ASHE-ERIC Higher Education Report No. T. Washington, D.C.: Association for the Study of Higher Education.
- Cross, P. K. (1986). Using assessment to improve instruction. Paper presented at the 47th ETS Invitational Conference.

- Cross, P. K. (1988). Feedback in the classroom: Making assessment matter. Paper presented at the AAHE Assessment Forum. Third National Conference on Assessment in Higher Education.
- Curry, W. & Hager, E. (1987). Assessing general education: Trenton State College. Student outcomes assessment: What institutions stand to gain, No. 59, 57-66.
- Dong, H. K., Sung, Y. H., & Dohm, T. E. (1985). The validity of the Ball Aptitude Battery (BAB): I. Relationship to high school academic success. Educational and Psychological Measurement, 45(3), 627-37.
- Doyle, D. P. & Hartle, T. W. (1985). Leadership in education: Governors, legislators and teachers. Phi Delta Kappan, 67(1), 21-27.
- Ebel, R. L. (1965). Measuring educational achievement. Englewood Cliffs: Prentice-Hall, Inc.
- Ebel, R. L. (1972). Essentials of educational measurement. Englewood Cliffs: Prentice-Hall, Inc.
- Ewell, P. (1985). Assessment: What's it all about? Change, 17(6), 32-36.
- Ewell, P. (1987a). Establishing a campus-based assessment program. New Directions for Higher Education, No. 59, 9-23.
- Ewell, P. (1987b). Assessment: Where are we? Change, 19(1), 23-28.
- Funk, S. G., Sturner, R. A., & Green, J. A. (1986). Preschool prediction of early school performance: Relationship of McCarthy scales of children's abilities prior to school entry to achievement

- in kindergarten, first, and second grades. The Journal of School Psychology, 24(2), 181-94.
- Gougeon, D. (1985). CEEB SAT mathematics scores and their correlation with college performance in math. Educational Research Quarterly, 9(2), 8-11.
- Hagerty, B. M. K. & Stark, J. S. (1989). Comparing educational accreditation standards in selected professional fields. Journal of Higher Education, 60(1), 1-20.
- Hakstian, A. R. & Woolsey, L. K. (1985). Validity studies using the Comprehensive Ability Battery (CAB): IV. Predicting achievement at the university level. Educational and Psychological Measurement, 45(2), 329-41.
- Halpern, D. F. (1987). Student outcomes assessment: Introduction and overview. New Directions for Higher Education, No. 59, 5-8.
- Hanna, G. S., Sonnenschein, J. L. & Lenke, J. M. (1983). The contribution of work-sample test items, student-reported past grades, and student-predicted grades in forecasting achievement in first-year algebra. Educational and Psychological Measurement, 43(1), 243-49.
- Hawley, W. D. (1985). False premises, false promises: The mythical character of public discourse about education. Phi Delta Kappan, 67(3), 183-187.
- Heller, S. (April 13, 1988). Miami-Dade College begins project to bolster teaching by evaluating new professors and rewarding

- classroom performance. Chronicle of Higher Education, 34(31), A12-A13, A18.
- Hintze, J. L. (1987). Number Cruncher Statistical System. Kaysville: Jerry L. Hintze.
- Hodgkinson, H. L. (1983). Guess who's coming to college: Your students in 1990. A Research Report from the State-Nation Information Network for Independent Higher Education.
- Impara, J. C. & Stoker, H. W. (1985). Determining reliability and validity of licensing examinations. In J. C. Fortune (Ed.), Understanding testing in occupational licensing (pp. 65-86). San Francisco: Jossey-Bass, Inc. Publishers.
- Integrity in the college curriculum. (1985, February 13). Chronicle of Higher Education, pp. 12-16, 18-22, 24, 26-30.
- Jacobi, M., Astin, A., & Ayala, F. (1987). College student outcomes assessment: A talent development perspective. (ASHE-ERIC Higher Education Report No. 7). Washington, DC: Association for the Study of Higher Education.
- Jaschik, S. (1985, September 18). Public universities try tests and surveys to measure what students learn. Chronicle of Higher Education, 31(3), 1, 16.
- Jaschik, S. (1985, October 2). As states weigh 'value-added' tests, Northeast Missouri offers model. Chronicle of Higher Education, 31(5), 21-22.

- Jones, R. F. & Thomae-Forgues, M. (1984). Validity of the MCAT in predicting performance in the first two years of medical school. Journal of Medical Education, 59(6), 455-64.
- Keller, J. E. (1970, May). Higher education objectives: Measures of performance and effectiveness. Ford Foundation Program for Research in University Administration.
- Krueger, D. W. and Heisserer, M. L. (1987). Assessment and involvement: Investments to enhance learning. New Directions for Higher Education, No. 59, 45-56.
- Lindeman, R. H. (1967). Educational measurement. Glenview: Scott, Foresman, and Company.
- Loacker, G., Cromwell, L., & O'Brien, K. (1986). Assessment in higher education: To serve the learner. Assessment in American Higher Education: Issues and Contexts. U. S. Department of Education: Office of Educational Research and Improvement, 47-61.
- Manning, T. E. (1988). Are the secretary's intentions honorable? Academe, 74(4), 12-15.
- Marchese, T. J. (1986). College: Raising a new vision. Change, 18(6), 10-17.
- Massoth, N. A. (1985). The McCarthy Scales of Children's Abilities as a predictor of achievement: A five-year follow-up. Psychology in the Schools, 22(1), 10-13.
- McMillan, J. H. (1988). Beyond value-added education. Journal of Higher Education, 59(5), 564-579.

- Mehrens, W. A. & Lehmann, I. J. (1984). Measurement and evaluation in education and psychology. New York: Holt, Rinehart and Winston.
- Morgan, R. G., Cornick, M. F., & Kauder, W. R. (1985). Predicting student success in Intermediate Accounting I. Journal of Education for Business, 61(2), 80-84.
- Mortimer, K. P. (1985). Sense and nonsense about excellence. NASPA Journal, 22(2), 10-18.
- National Commission on Excellence in Education. (1983). A nation at risk. Washington, DC: U.S. Government Printing Office. (ERIC Document Reproduction Service No. ED 226 006).
- National Governors' Association. (1986). Time for results: The governors' 1991 report on education. Washington, D.C.: National Governors' Association Center for Policy Research and Analysis.
- Osigweh, Yg., C. A. B. (1986). An evaluation model of training outcomes for higher education. Educational Evaluation and Policy Analysis, 8(2), 167-178.
- Pace, C. R. (1984a). Historical perspectives on student outcomes: Assessment with implications for the future. NASPA Journal, 22(2), 10-18.
- Pace, C. R. (1984b). Measuring the quality of college student experiences. Los Angeles: University of California, Higher Education Research Institute, Graduate School of Education.
- Pascarella, E. T. (1986). Are value-added analyses valuable? Paper presented at the 47th Educational Testing Service Invitational Conference, New York City.

- Payne, D. A. (1974). The assessment of learning: Cognitive and affective. Lexington: D.C. Heath and Company.
- Phillips, R. C. (1968). Evaluation in education. Columbus: Charles E. Merrill Publishing Company.
- Prevatt, M. B. (1984). The development of a clothing construction exemption test. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Reilly, T. P., Drudge, O. W., Rosen, J. C., Loew, D. e., & Fischer, M. (1985). Concurrent and predictive validity of the WISC-R, McCarthy Scales, Woodcock-Johnson, and academic achievement. Psychology in the Schools, 22(4), 380-82.
- Richmond, J. (1986). The importance of student involvement: A dialogue with Alexander Astin. Journal of Counseling and Development, 65, 92-95.
- Schwarz, P. A. (1971). Prediction instruments for educational outcomes. In R. L. Thorndike (Ed.), Educational measurement (pp. 303-331). Washington, DC: American Council on Education.
- Smith, H. L. (1986). Testimony for the national governors association task force on college quality on behalf of the American Association of State Colleges and Universities. Paper collected as part of the American Association for Higher Education Assessment Forum.
- Southern Association of Colleges and Schools. (1989-1990). Criteria for accreditation: Commission on colleges. (Available from [SACS, 1866 Southern Lane, Decatur, GA 30033-4097]).

- Spanghel, S. D. (1987). The push to assess: Why it's feared and how to respond. Change, 19(1), 35-39.
- Stanley, J. C. & Hopkins, K. D. (1972). Educational and psychological measurement and evaluation. Englewood Cliffs: Prentice-Hall, Inc.
- Szymczuk, M. & Frerich, D. (1985). Using standardized tests to predict achievement in an introductory high school computer course. AEDS Journal, 19(1), 20-27.
- Tan, D. L. (1986). The assessment of quality in higher education: A critical review of the literature and research. Research in Higher Education, 24(3), 233-265.
- Taylor, T. (1985). A value-added student assessment model: Northeast Missouri State University. Economics of Education Review, 4(4), 341-350.
- Terenzini, P. T. (1989). Assessment with open eyes: Pitfalls in studying student outcomes. Journal of Higher Education, 644-664.
- Text of new report on excellence in undergraduate education. (1984, October 24). Chronicle of Higher Education, pp. 35-37, 38, 40-49.
- Text of Secretary Bennett's address last week at a Harvard University anniversary celebration. (1986, October 15). Chronicle of Higher Education, p. 27-30.
- Thornell, J. G. & McCoy, A. (1985). The predictive validity of the GRE for subgroups of students in different academic disciplines. Educational and Psychological Measurement, 45(2), 415-19.
- To Reclaim a Legacy: Text of report on humanities in education. (1984, November 28). The Chronicle of Higher Education, p. 16-21.

Turnbull, W. W. (1985). Are they learning anything in college?

Change, 17(6), p. 23.

Vandament, W. E. (1987). A state university perspective on student outcomes assessment. New Directions for Higher Education, No. 59, 25-58.

Virginia Polytechnic Institute and State University Plan for Assessing Institutional Effectiveness: Undergraduate Student Outcomes. (1987). (Available from the Office of the University Provost, 201 Burruss Hall, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061).

Virginia Tech Alumni Association 1986-87 Annual Report. (1987).

(Available from Virginia Tech Alumni Association, Alumni Hall, Virginia Tech, Blacksburg, VA 24061).

Warren, J. (1987). Assessment at the source. Liberal Education, 73(3), p. 2-6.

APPENDIX A

Pretest Survey

Please complete the information in this packet to the best of your ability. Do not discuss your answers with any of your classmates. The information you provide is confidential and will not affect the grade you receive in this class. You may write on this page, however for the test itself, please use the opscan sheet if one has been provided.

1. Student ID Number _____
2. Classification
 - a. Freshman b. Sophomore c. Junior d. Senior
3. Are you majoring in the Department of Clothing and Textiles?
 - a. yes b. no
4. If you answered "yes" to #3 which option have you selected?
 - a. Apparel Design
 - b. Fashion Merchandising
 - c. Fashion Merchandising and Marketing Education
 - d. Textiles
 - e. Both a & b
 - f. Undecided
5. If you answered "no" to #2, what is your major? _____
6. Check which of the following courses you have taken or are currently taking in the Department of Clothing & Textiles (mark all that apply).

	Have Taken	Currently Taking
a. CT 1004 Clothing & People	_____	_____
b. CT 2404 Basic Apparel Assembly	_____	_____
c. CT 2414 Basic Apparel Assembly Lab	_____	_____
d. CT 2604 Intro to Fashion Industry	_____	_____
e. CT 3204 Textile Science	_____	_____
7. What is your reason for taking this class you are currently in?
 - a. Required course in my major
 - b. Required course outside my major
 - c. Elective in my major
 - d. Elective outside my major
 - e. None of the above - Please specify _____
8. Are you taking this course for
 - a. Pass Fail? b. A-F? c. Audit?
9. What grade do you plan to earn in this course?
 - a. A or A- b. B+, B, or B-
 - c. C+, C, or C- d. D+, D, or D-
 - e. None of the above

Posttest Survey

Please complete the information in this packet to the best of your ability. Do not discuss your answers with any of your classmates. The information you provide is confidential and will not affect the grade you receive in this class. You may write on this page, however for the test itself, please use the opscan sheet.

1. Student ID Number _____
2. Classification
 - a. Freshman
 - b. Sophomore
 - c. Junior
 - d. Senior
3. Are you a major in the Department of Clothing and Textiles?
 - a. yes
 - b. no
4. If you answered "yes" to #3 which option have you selected?
 - a. Apparel Design
 - b. Fashion Merchandising
 - c. Fashion Merchandising and Marketing Education
 - d. Textiles
 - e. Both a & b
 - f. Undecided
- 4b. If you answered "yes" to #3 did you transfer into the department?
 - a. yes
 - b. no
5. Check which of the following courses you have taken or are currently taking in the Department of Clothing & Textiles (mark all that apply).

	Have Taken	Currently Taking
a. CT 1004 Clothing & People	_____	_____
b. CT 2404 Basic Apparel Assembly	_____	_____
c. CT 2414 Basic Apparel Assembly Lab	_____	_____
d. CT 2604 Intro to Fashion Industry	_____	_____
e. CT 3204 Textile Science	_____	_____
6. How many out-of class organizations have you been involved in during your college attendance? _____
7. Estimate the number of hours you spend per week on school related activities (i.e., studying, group sessions, library work, etc.).
 - a. 20 hours or less
 - b. 21-29 hours
 - c. 30 - 39 hours
 - d. 40 hours or more
8. Have you been involved in paid employment for 10 or more hours per week during your college attendance?
 - a. yes
 - b. no
9. If yes, was (or is) this employment related to your major area of study?
 - a. yes
 - b. no

Department of Clothing and Textiles
Core Knowledge Exam

Directions: For the following statements select the answer that best completes each statement. Place all answers on the answer sheet provided.

10. In periods of greater social and economic freedom for women, their dress tends to be
 - a. softer and more feminine.
 - b. more characteristic of male attire.
 - c. more variable, following fewer traditional conventions.
11. In every society the type of dress worn is
 - a. influenced by accepted standards of right and wrong.
 - b. completely left to individual tastes and preferences.
 - c. based on a person's social class or status.
12. There are strong social pressures influencing dress even where there are no specific laws or rules.
 - a. True
 - b. False
13. Despite the cyclical nature of fashion, a fashion never returns in the same form as it was.
 - a. True
 - b. False
14. Laws regulating amount and type of clothing worn are called
 - a. dress codes.
 - b. sumptuary legislation
 - c. mores.
 - d. quotas.
15. Modesty is
 - a. the best explanation for the origin of clothing.
 - b. conceptualized differently from culture to culture.
 - c. dependent on how much clothing one wears.
 - d. unfashionable in our society.
 - e. all of the above.
16. First impressions are based on
 - a. intelligence, dress, and economics.
 - b. behavior, intelligence, and dress
 - c. posture, behavior, and dress.
 - d. politics, behavior, and intelligence.

17. High fashion refers to clothing items that
 - a. have reached the highest point of acceptance on the normal fashion curve.
 - b. are found only at the beginning of the normal fashion cycle.
 - c. are accepted by the highest (largest) number of people.
 - d. have sudden bursts of popularity among a large group of people.
18. Clothing is a useful tool with which to study culture because
 - a. it is obvious.
 - b. all peoples use it in some form.
 - c. it is found in the tombs and burial sites of many ancient peoples.
 - d. all of the above.
 - e. only a and b.
19. In a democracy people are free to dress any way they choose.
 - a. True
 - b. False
20. Trade Names
 - a. are defined by the Federal Trade Commission according to the chemical composition of the fiber.
 - b. selected by the fiber manufacture.
 - c. are required by law on all apparel labels.
 - d. have several types in each family of fibers.
21. The major governmental agency concerned with enforcing consumer protection legislation is the
 - a. Federal Trade Commission.
 - b. The National Bureau of Standards.
 - c. Department of Commerce.
 - d. The Federal Bureau of Investigation.
 - e. Consumer's Union.
22. Exempt from the Federal Trade Commission's Permanent Care Labeling Ruling are
 - a. foot and hand coverings.
 - b. hats and other head coverings including scarves.
 - c. handkerchiefs and sheer blouses.
 - d. all of the above.
 - e. only 1 and 2.

23. The Textile Fiber Products Identification Act does not require labels to include
- a. the name of the country of origin if imported.
 - b. the generic name of all fibers present in the garment.
 - c. the name of the manufacturer of the product.
 - d. the percentage of fibers present in the garment.
 - e. all of the above are required.
24. Flammability of textile products is related to
- a. fiber content.
 - b. fabric and yarn construction.
 - c. garment construction.
 - d. fabric finishes.
 - e. all of the above.
25. "Style" when referring to clothing
- a. is synonymous with fashion.
 - b. changes with fashion changes.
 - c. is part of the current norms.
 - d. is unchanging over long periods of time.
26. Fashion has been criticized because it
- a. may be detrimental to health.
 - b. can be uncomfortable.
 - c. is economically wasteful.
 - d. all of the above.
27. The study of clothing is an interdisciplinary study. This means that
- a. the subject matter involved depends upon a well disciplined approach.
 - b. one must relate, in an orderly fashion, the clothing behavior of the past to that of the present.
 - c. many subject matter areas contribute to the study of clothing.
28. It can be demonstrated through systematic analysis that clothing behavior is not the random and purposeless activity it is often presumed to be.
- a. True
 - b. False
29. "Women's Wear Daily" is a
- a. consumer fashion magazine.
 - b. consumer daily newspaper.
 - c. trade publication.
 - d. none of the above.

30. In the early stages of an apparel industry in any developing country the quality of mass produced clothing is
- generally poor.
 - better in children's apparel than in clothing for adults.
 - clearly defined by management for the customer
 - comparable to clothing quality in developed countries
31. Automation has been used in the fabric producing textile industry more than in the apparel industry because
- the fabric producers have greater profit margins and can afford automated equipment.
 - the fabric is less limp during production than when cut in garment parts for apparel.
 - it is easier to handle fabric with automated equipment while it is in continuous yardage than after it has been cut into garment pieces.
 - the production processes in the apparel industry have not needed to be automated.
32. The quality of the apparel product is affected
- by each stage of production.
 - by the balance of production and on time delivery.
 - by the retail store in which it is sold.
 - solely by the price of the product.
33. One of the broadest movements in the U.S. in the 1980's to save jobs in the domestic textile and apparel industries was
- the organization of Textile/Clothing Technology Corporation.
 - Quick Response.
 - emphasis on quality.
 - Crafted with Pride in the U.S.A.
34. A relative new synthetic fiber that has been used extensively in men's undergarments and in active wear is
- polypropylene.
 - polyester.
 - triacetate.
 - rayon.
35. The 4-Point System for grading fabrics
- is a mandatory grading system used to prevent the buyer from receiving piece goods with any flaws in the fabric.
 - is a voluntary grading system that allows for some fabric flaws to be included in a shipment of goods to the apparel manufacturer.
 - is a voluntary grading system in which the fabric manufacturers establish the acceptable tolerance level for fabric flaws.
 - is a voluntary system used to prevent the buyer from receiving piece goods with any flaws in the fabric.

36. Apparel items that are exempt from the Permanent Care Labeling Ruling are
 - a. hats, gloves, and slippers.
 - b. hats, mittens, and half slips.
 - c. gloves, scarves, and pantyhose.
 - d. gloves, pantyhose, and half slips.
37. The relationship of distances, areas, and/or parts of a garment is based on the design principle of
 - a. scale.
 - b. harmony.
 - c. balance.
 - d. proportion.
38. Mandatory labeling established by the Federal Trade Commission in 1972 is the
 - a. Textile Fiber Identification Act.
 - b. Permanent Care Labels Ruling.
 - c. Wool Products Labeling Act.
 - d. Flammable Fabrics Act.
39. The crossgrain yarn in knitted fabrics is called
 - a. weft.
 - b. wale.
 - c. course.
 - d. warp.
40. Which of the following is not an example of a plain seam?
 - a. French seam
 - b. Princess seam
 - c. Corded seam
 - d. Bias seam
41. A heat sealed seam might be used in
 - a. nylon fabric.
 - b. acetate fabric.
 - c. rayon fabric.
 - d. polyester/cotton blends.
42. Seams should be finished according to the
 - a. dictates of the seam type, fabric, and projected garment care.
 - b. most economical method.
 - c. fabric construction and the width of the seam allowance.
 - d. the preference of the manufacturer.

43. The organizations that develop and monitor the majority of textile testing methods are
- International Standards Organization and the American Society of Testing Materials.
 - American Society of Testing and Materials and the U.S. Government.
 - Association of Textile Chemists and Colorists and the International Standards Organization.
 - Association of Textile Chemists and Colorists and the American Society of Testing and Material.
44. Neckline treatments may consist of bindings, facings, or bands. Bindings will fit a curved neckline better when
- cut the same shape as the neckline.
 - cut on the bias grain.
 - cut on the lengthwise grain.
 - cut on the crosswise grain.
45. A characteristic of a flat collar is
- the neckline edge is nearly identical to the neckline of the garment.
 - the neckline edge is straighter than the neckline of the garment.
 - the outer edge always takes the shape of a sailor collar.
 - always a Peter Pan collar.
46. The sleeve most difficult to sew into a garment is a
- low-cap set-in.
 - raglan.
 - high-cap set-in.
 - kimono.
47. A gusset is used in the underarm area to provide extra room for body movement in a
- high-cap sleeve.
 - fitted kimono sleeve.
 - loose kimono sleeve.
 - raglan sleeve.
48. If you design a circular skirt, you should specify a hem width of
- 1" or less to eliminate bulk in the hem.
 - at least 2" to add more weight around the bottom of the skirt.
 - 1" or less to reduce production costs.
 - 1" for budget priced skirt, 2" for a popular priced skirt, or 3" for a better priced skirt.

49. The understructure fabric that is attached to the back of each garment section before the pieces are sewn together is called
- interlining.
 - lining.
 - interfacing.
 - underlining.
50. The standard width of a seam on a standard commercial pattern is
- $\frac{3}{8}$ ".
 - $\frac{1}{2}$ ".
 - $\frac{5}{8}$ ".
 - $\frac{3}{4}$ ".
51. Understitching is recommended for collars and neck facings in order to
- prevent the outer seam from rolling to the right side.
 - reinforce the outer seam.
 - prevent the outer seam allowance from raveling.
 - all of the above are correct.
52. As a general rule, the weight of an interfacing fabric should be
- much lighter than the fabric with which it is used.
 - slightly lighter than the fabric with which it is used.
 - near the same weight as the fabric with which it is used.
 - slightly heavier than the fabric with which it is used.
53. Napped fabric
- often require additional fabric.
 - require using a one-way layout.
 - include fabrics such as corduroy, flannel, and knits.
 - all of the above.
54. In comparing the pattern pieces for the armhole (or armhole) and the sleeve cap of a fitted set-in sleeve, one would find that
- the sleeve cap is larger than the armhole into which it will be placed.
 - the sleeve cap is smaller than the armhole into which it will be placed.
 - armhole and sleeve cap are equal in size.
55. The stitch most often used to secure fasteners (hooks, eyes, and snaps) is the
- catch stitch.
 - blind stitch.
 - buttonhole or blanket stitch.
 - slip stitch.

56. Threads running parallel to the selvage are referred to as
- crosswise grain.
 - lengthwise grain.
 - bias grain.
 - warp yarns.
 - both b and d are correct.
57. The zipper application where both sides of the fabric lap towards the center of the zipper is called the
- lapped application.
 - centered application.
 - exposed zipper.
 - invisible zipper.
58. The fabric design or weave which would require the most material for a layout is
- even plaids.
 - uneven plaids.
 - small gingham checks.
 - napped fabric.
59. The direction in a woven fabric with the greatest stretch is
- garment bias.
 - crosswise grain.
 - lengthwise grain.
 - true bias.
 - all of the above
60. The stitch length control lever (or dial) on the sewing machine
- controls the width of the zig-zag.
 - controls the distance the feed dog moves the fabric.
 - regulates the speed of the machine.
 - none of the above.
61. Inconspicuous topstitching done in the crevice of a previously stitched seam is called
- understitching.
 - staystitching.
 - stitching in the ditch.
 - topstitching.
62. Horizontal darts in a garment are usually pressed
- in either direction.
 - open.
 - up.
 - down.

63. The more circular the skirt becomes the narrower the hem should be.
- a. True
 - b. False
64. "Clean finishing" means
- a. cutting and tying all loose threads on inside of garment.
 - b. pressing the garment on the wrong side to improve appearance.
 - c. turning under and stitching all raw edges on inside of garment.
 - d. finishing the outer edge of a facing by turning under the fabric 1/4" and stitching close to folded edge.
65. Interfacing is designed to
- a. prevent stretch and provide shape.
 - b. protect seams.
 - c. give warmth.
 - d. prevent wrinkles.
66. A size 16 regular machine needle would be appropriate for use with
- a. wide-wale corduroy.
 - b. polyester crepe.
 - c. cotton double-knit.
 - d. upholstery fabric
67. The main purpose of staystitching is
- a. to mark seam lines.
 - b. to prevent fabric from raveling.
 - c. to prevent fabric from stretching.
 - d. to allow fabric to be gathered.
68. The measurement which is the key to one's pattern size for dresses, suits, and coats is
- a. hips.
 - b. waist.
 - c. bust.
 - d. back waist length.
69. Generally, the collar should be attached to a garment
- a. before the neckline facing is attached.
 - b. after the sleeves are sewn in.
 - c. after the neckline facing is attached.

70. Which of the following is the most direct method of determining customer fashion preferences for purposes of fashion prediction?
- Studying selling records
 - Consulting fashion suppliers
 - Exchanging information with noncompetitors
 - Subscribing to fashion periodicals
71. The prevailing or accepted style in dress or personal decoration established or adopted during a particular time or season is a definition of
- a fad.
 - fashion.
 - a design.
 - classics.
72. The trickle-down theory of fashion leadership explains why people dress like
- movie-stars.
 - the young.
 - their peers.
 - the poor.
73. The part of the fashion industry that produces trims, threads, zippers, and interfacing
- primary suppliers.
 - secondary suppliers.
 - accessory suppliers.
 - converter.
74. The three main types of silhouettes seen in fashion throughout history are
- back-fullness, bell-shaped, triangular.
 - tubular, round, bell-shaped.
 - tubular, bell-shaped, back-fullness.
 - tubular, rectangular, bell-shaped.
75. _____ is a firm that buys or handles the greige good from mills and contracts them out to finishing plants to have them finished.
- converter
 - contractor
 - tanner
 - jobber
76. What determines whether a style becomes a fashion is
- promotion and publicity.
 - customer acceptance.
 - designer creativity.
 - irrational factors.

77. West Germany is famous for its
- workmanship and fabrics.
 - low-wage production centers.
 - innovative trendy designs.
 - international textile and apparel fairs.
78. The work of top Japanese fashion designers is
- intended to be mass merchandised.
 - based on traditional Western concepts.
 - focused on new ways of constructing clothes.
 - marketed almost exclusively in Japan.
79. Which of the following is an illegal method of getting around quotas?
- Transferring production to countries with unused quotas
 - Authorizing production in countries with excess quotas
 - Producing in developing countries
 - Transshipping through countries with quota availability
80. To curtail evasion, new country-of-origin laws in the United States have determined that a garment originates where
- the first manufacturing steps are taken.
 - it is assembled.
 - shipment to the United States is made.
 - the last finishing steps are performed.
81. Section 807 of the tariff schedule encourages foreign
- contracting.
 - trade barriers.
 - investment.
 - price increases.
82. Small retailers are less likely to take advantage of
- foreign trade shows in the United States.
 - foreign producer showrooms in the United States.
 - store-owned foreign buying offices.
 - foreign commissionaires and agents.
83. Which of the following methods of penetrating foreign markets is used least?
- Licensing agreements
 - Direct exports
 - Direct ownership of factories
 - Joint-ownership ventures

84. Which of the following is a fashion information service?
- Youth Fashion Guild
 - Yankelovich, Skelly, and White
 - The Fashion Group
 - IM International
85. The Daily News Record is a trade publication for
- menswear and textiles.
 - children's wear.
 - accessories.
 - haute couture.
86. The International Ladies Garment Worker's Union gained strength as an immediate result of the
- rise of sweatshops.
 - Triangle Shirtwaist fire.
 - publication of Godey's Lady's Book.
 - World War I.
87. When a garment is knocked off, the
- design has been copied or adopted.
 - price has been marked down.
 - style has been copyrighted.
 - licensee has paid a royalty fee.
88. The factor which contributed most to making New York the center of the garment industry was the
- invention of the sewing machine.
 - arrival of waves of immigrants eager to work.
 - availability of buildings on Seventh Avenue.
 - decentralization of production.
89. _____ is the amount by which the value of exports exceed imports.
- Quotas
 - Trade deficit
 - Tariff
 - Factoring
90. The care label must be permanently attached for the life of the article.
- True
 - False
91. In the care of a garment, a care label must provide washing and drycleaning instructions.
- True
 - False

92. When protein fibers and cellulosic fibers are compared:
- a. protein fibers are weaker when wet.
 - b. cellulosic fibers have greater elongation and recovery.
 - c. cellulosic fibers are stronger.
 - d. both protein and cellulosic fibers have the same burning behavior.
93. When natural and man-made fibers are compared _____ fibers are chemically modified by 15% weight through copolymerization to achieve specific properties for end use suitability.
- a. natural
 - b. man-made
94. Which of the following fabrics has a rib weave and can be used for outer-wear to resist wind and water penetration?
- a. Poplin
 - b. Brocade
 - c. Percale
 - d. Damask
 - e. Corduroy
95. Raschel is a fabric using which of the following textile constructions?
- a. Woven
 - b. Knit
 - c. Nonwoven
 - d. Laminated or composite
 - d. None of the above
96. Which of the following fabrics has a complex or novelty weave with a combination of plain, twill, satin weaves, and a raised pattern?
- a. Poplin
 - b. Bengaline
 - c. Brocade
 - d. Raschel
 - e. Damask
97. Which of the following descriptions corresponds with a needlepunch nonwoven construction?
- a. Interlooped or interlaced thread provides stability
 - b. Mechanical binding of fiber webs
 - c. Filament fibers bound by adhesive

98. Which of the following dye classes may require processing with carriers?
- a. Disperse
 - b. Azoic
 - c. Sulfur
 - d. Direct
 - e. None of the above
99. Which of the following fabrics has a weft pile woven construction with pile stacked in vertical columns referred to as "wales"?
- a. Poplin
 - b. Marquisette
 - c. Velveteen
 - d. Velvet
 - e. Corduroy
100. Which of the following dye classes has a chemical reaction with the textile fibers?
- a. Reactive
 - b. Azoic
 - c. Disperse
 - d. Acid
 - e. None of the above
101. Which of the following fiber types could be used in a garment designed to meet the Federal Trade Commission flammability standards for children's sleepwear?
- a. Cotton
 - b. Flax
 - c. Rayon
 - d. Modacrylic
102. Which of the following fibers is the weakest fiber, especially when wet, but has good elongation and recovery, and high moisture regain?
- a. Wool
 - b. Nylon
 - c. Acrylic
 - d. Modacrylic
 - e. Rayon
103. Which of the following fibers is stiff, will not burn, and requires special handling?
- a. Polyester
 - b. Olefin
 - c. Silk
 - d. Aramid
 - e. Glass

104. Which of the following fibers is appropriate for apparel, household, and industrial uses and has the following properties: not acid resistant, not sunlight resistant, melts with burning?
- Rayon
 - Polyester
 - Nylon
 - Aramid
 - Olefin
105. Which of the following fibers is lighter than water and does not absorb water, is resistant to acids and bases, but can be attacked by some organic cleaning solutions?
- Spandex
 - Olefin
 - Nylon
 - Aramid
 - Silk
106. Which man-made fiber has poor dimensional stability and is weaker when wet?
- Polyester
 - Nylon
 - Rayon
 - Wool
 - Olefin
107. A fabric which has been treated with a chemical that increases the strength and luster in cotton and linen fabrics has been
- calendered.
 - mercerized.
 - embossed.
 - moired.
 - parchmentized.
108. A process which smooths and straightens fabric when it is passed between heated rollers is
- moireing.
 - embossing.
 - mercerization.
 - calendering.
 - parchmentization.
109. A finishing process in which two layers of ribbed fabric are pressed together off grain is
- parchmentization.
 - calendering.
 - embossing.
 - mercerization
 - moireing.

APPENDIX B

Protocol for Subtest Sample

1. Which core course(s) were you enrolled in last semester?
2. Did you find the class(es) interesting? If so, from what perspective, the information taught or teacher performance, or both? Could you elaborate?
3. Could you estimate the number of times you missed class?
4. Did you get a syllabus?

Were these objectives on the syllabus?

Was there an outline of the course on the syllabus?

How did you use the syllabus during the semester? (i.e., do you review the syllabus for policy information, to check progress, in preparation for a test, etc.)

5. Review these objectives. In your opinion, were all of these objectives covered during the course of the semester?

Which objectives seemed to have been given the most attention?

6. Do you feel you achieved these learning objectives?

7. Review the test items? Was all of the information (stressed) taught during the course of the semester?

If not, which items were stressed more than others or not taught?

8. How often did you review your class notes and read your text?
Possible options:

Frequently (Immediately after class)

Only before a test

Occasionally

Prior to each class period in preparation for upcoming class

9. Has any of the information taught in this core course been covered in other core courses or other courses you have taken?

Is there a building of information from the first time it is introduced through the other courses?

Protocol for Instructors

1. Which core course(s) did you teach last semester?
2. How many times have you taught the course(s) including last semester?
3. Review the test items? Of the items on the test, which items were covered in the most detail or given the most emphasis during the semester?
4. Are there any items that were not covered at all during the semester? Please identify those items.
5. Was your final examination comprehensive?
6. What was class attendance like?

On the average, what would you say was the attendance trend?

Do you feel this attendance was reflected in the grades students received?
7. How often do you refer to the course outline?

Do you refer to your course outline during the course of the semester to check progress (i.e., whether you are covering the information)?
8. Review the objectives. In your opinion, which of the objectives were covered most adequately and took up the bulk of the semester?

Do you regularly look back on objectives to check whether students are meeting course objectives?

How often do you refer students to the syllabus to check for policy or topics covered?

How do you use the objectives on the syllabus?
As a guide to organize the course?
As a guide in grading student work?
As a guide in developing tests?
Other?
9. How often do you feel you have to cut off information at the end of the semester, because of the lack of time?

Did this happen last semester?
10. In general how did you feel about the class as a whole? (i.e., student participation and responsiveness)

Protocol for Seniors

1. What is your option area?
2. Have you taken all core courses in the department?
3. Has any of the information taught in the core courses been covered in other courses in the department or outside of the department?
Please elaborate.

Is there a building of information from the first time it is introduced through your senior year?

4. Discuss the courses? Did you find them interesting? From what perspective?
5. On the average, how often did you miss class?
6. Do you recall getting a syllabus in all department courses?
How did you use the syllabus?
7. How often did you review your class notes and read the text in your classes? You may be specific or generalize to all courses.
8. Review the test items. Identify the items you feel were not covered in the core courses and those that were heavily emphasized?
9. How do you feel about your success in department courses (particularly the core courses)?
10. What are your career plans?
Do you feel our program has adequately prepared you for your plans?
11. What are your suggestions for improving the courses or the program as a whole?
12. Do you feel there should be some measure of your achievement (other than course grades or QCA) in the curriculum?

APPENDIX C

Course Objectives and Related Subtest Items for Core Courses

CT 1004 - Clothing and People

1. Appreciate the relationships between clothing and textiles and the culture in which we live.
 - (Item 11) In every society the type of dress worn is [more variable, following fewer traditional conventions].
 - (Item 18) Clothing is a useful tool with which to study culture because [it is obvious, all peoples use it in some form, and it is found in the tombs and burial sites of many ancient peoples].
2. Understand the importance of dress as a means of self expression and a mode of communication among individuals, groups, and cultures.
 - (Item 16) First impressions are based on [posture, behavior, and dress].
 - (Item 17) High fashion refers to clothing items that [are found only at the beginning of the normal fashion cycle].
 - (Item 25) "Style" when referring to clothing [is unchanging over long periods of time].
 - (Item 26) Fashion has been criticized because it [may be detrimental to health, can be uncomfortable, and is economically wasteful].
3. Understand the forces that influence clothing behavior.
 - (Item 10) In periods of greater social and economic freedom for women, their dress tends to be [more variable, following fewer traditional conventions].
 - (Item 12) There are strong social pressures influencing dress even where there are no specific laws or rules [true].
 - (Item 13) Despite the cyclical nature of fashion, a fashion never returns in the same form as it was. [True]
 - (Item 14) Laws regulating amount and type of clothing worn are called [sumptuary legislation].

- (Item 15) Modesty is [conceptualized differently from culture to culture].
 - (Item 19) In a democracy people are free to dress any way they choose [false].
 - (Item 28) It can be demonstrated through systematic analysis that clothing behavior is not the random and purposeless activity it is often presumed to be [true].
4. Recognize the components of apparel-fibers, fabrics, and finishes which help individuals make clothing choices that provide maximum satisfaction.

(Related items are in CT 3204 subtest.)

5. Increase satisfaction in personal use of clothing.

- (Item 20) Trade Names [are required by law on all apparel labels].
- (Item 21) The major governmental agency concerned with enforcing consumer protection legislation is the [Federal Trade Commission].
- (Item 22) Exempt from the Federal Trade Commission's Permanent Care Labeling Ruling are [foot and hand coverings, hats and other head coverings including scarves, and handkerchiefs and sheer blouses].
- (Item 23) The Textile Fiber Products Identification Act does not require labels to include [the name of the country of origin if imported, the generic name of all fibers present in the garment, the name of the manufacturer of the product, and the percentage of fibers present in the garment].
- (Item 24) Flammability of textile products is related to [fiber content, fabric and yarn construction, garment construction, and fabric finishes].

CT 2404 - Basic Apparel Assembly Lecture

1. Interpret terminology associated with the manufacturing, merchandising, and consumer use of apparel products.
- (Item 14) A relatively new synthetic fiber that has been used extensively in men's undergarments and in active wear is [polypropylene].

- (Item 15) The 4-Point System for grading fabrics [is a voluntary grading system that allows for some fabric flaws to be included in a shipment of goods to the apparel manufacturer].
- (Item 16) Apparel items that are exempt from the Permanent Care Labeling Ruling are [hats, gloves, and slippers].
- (Item 18) Mandatory labeling established by the Federal Trade Commission in 1972 is the [Permanent Care Labels Ruling].
- (Item 19) The crossgrain yarn in knitted fabrics is called [course].
- (Item 20) Which of the following is not an example of a plain seam? [French seam].
- (Item 21) A heat sealed seam might be used in [nylon fabric].
- (Item 23) The organizations that develop and monitor the majority of textile testing methods are [Association of Textile Chemists and Colorists and the American Society of Testing and Material].
- (Item 25) A characteristic of a flat collar is [the neckline edge is nearly identical to the neckline of the garment].
- (Item 26) The sleeve most difficult to sew into a garment is a [high cap set-in].
- (Item 27) A gusset is used in the underarm area to provide extra room for body movement in a [fitted kimono sleeve].
- (Item 29) The understructure fabric that is attached to the back of each garment section before the pieces are sewn together is called [underlining].

2. Understand basic concepts and principles that govern production processes.

- (Item 11) Automation has been used in the fabric producing textile industry more than in the apparel industry because [it is easier to handle fabric with automated equipment while it is in continuous yardage than after it has been cut into garment pieces].

3. Understand the importance of efficiently managing time, equipment, and supplies in the garment assembly process.

(Item 13) One of the broadest movements in the U.S. in the 1980's to save jobs in the domestic textile and apparel industries was [Crafted with Pride in the U.S.A.].

4. Understand elements in the assessment of apparel quality.

(Item 10) In the early stages of an apparel industry in any developing country the quality of mass produced clothing is [generally poor].

(Item 12) The quality of the apparel product is affected [by each stage of production].

(Item 17) The relationship of distances, areas, and/or parts of a garment is based on the design principle of [proportion].

(Item 22) Seams should be finished according to the [dictates of the seam type, fabric, and projected garment care].

(Item 24) Neckline treatments may consist of bindings, facings, or bands. Bindings will fit a curved neckline better when [cut on the bias grain].

(Item 28) If you design a circular skirt, you should specify a hem width of [1" or less to eliminate bulk in the hem].

CT 2414 - Basic Apparel Assembly Laboratory

1. Understand and apply terminology used in the apparel process.

(Item 10) The standard width of a seam on a standard commercial pattern is [5/8"].

(Item 13) Napped fabric [often require additional fabric, require using a one-way layout, and include fabrics such as corduroy, flannel, and knits].

(Item 16) Threads running parallel to the selvage are referred to as [lengthwise grain, warp yarns].

(Item 17) The zipper application where both sides of the fabric lap towards the center of the zipper is called the [centered application].

(Item 18) The fabric design or weave which would require the most material for a layout is [uneven plaids].

(Item 19) The direction in a woven fabric with the greatest stretch is [true bias].

(Item 25) "Clean finishing" means [finishing the outer edge of a facing by turning under the fabric 1/4" and stitching close to folded edge].

2. Utilize concepts, principles, and techniques of garment assembly in the production of finished consumer products.

(Item 11) Understitching is recommended for collars and neck facings in order to [prevent the outer seam from rolling to the right side].

(Item 12) As a general rule, the weight of an interfacing fabric should be [slightly lighter than the fabric with which it is used].

(Item 14) In comparing the pattern pieces for the armhole (or armhole) and the sleeve cap of a fitted set-in sleeve, one would find that [the sleeve cap is larger than the armhole into which it will be placed].

(Item 15) The stitch most often used to secure fasteners (hooks, eyes, and snaps) is the [buttonhole or blanket stitch].

(Item 21) Inconspicuous topstitching done in the crevice of a previously stitched seam is called [stitching in the ditch].

(Item 22) Horizontal darts in a garment are usually pressed [down].

(Item 23) The more circular the skirt becomes the narrower the hem should be. [True]

(Item 25) Interfacing is designed to [prevent stretch and provide shape].

(Item 27) The main purpose of staystitching is [to prevent fabric from stretching].

(Item 29) Generally, the collar should be attached to a garment [before the neckline facing is attached].

3. Demonstrate through garment construction the importance of time, equipment, supplies, and fabric in evaluating garment quality.

(Item 20) The stitch length control lever (or dial) on the sewing machine [controls the distance the feed dog moves the fabric].

(Item 26) A size 16 regular machine needle would be appropriate for use with [upholstery fabric].

4. Apply some basic fitting principles in the construction of garments.

(Item 28) The measurement which is the key to one's pattern size for dresses, suits, and coats is [bust].

CT 2604 - Introduction to the Fashion Industry

1. To acquire a basic understanding of the nature of fashion.

(Item 10) Which of the following is the most direct method of determining customer fashion preferences for purposes of fashion prediction? [studying selling records]

(Item 11) The prevailing or accepted style in dress or personal decoration established or adopted during a particular time or season is a definition of [fashion].

(Item 12) The trickle-down theory of fashion leadership explains why people dress like [movie-stars].

(Item 14) The three main types of silhouettes seen in fashion throughout history are [tubular, bell-shaped, back-fullness].

(Item 16) What determines whether a style becomes a fashion is [customer acceptance].

(Item 27) When a garment is knocked off, the [design has been copied or adopted].

2. To analyze the economic, sociological, and psychological factors affecting contemporary fashion; to become acquainted with literature and resource materials dealing with these factors.

(Item 25) The Daily News Record is a trade publication for [menswear and textiles].

3. To understand the history and development of the various segments of the fashion industry and the unique and complete interdependency of these segments.

(Item 24) Which of the following is a fashion information service? [The Fashion Group]

- (Item 26) The International Ladies Garment Worker's Union gained strength as an immediate result of the [Triangle Shirtwaist fire].
- (Item 28) The factor which contributed most to making New York the center of the garment industry was the [arrival of waves of immigrants eager to work].
4. To gain an understanding and appreciation of the purposes served by foreign and domestic markets and to examine the role of designers in these markets.
- (Item 17) West Germany is famous for its [international textile and apparel fairs].
- (Item 18) The work of top Japanese fashion designers is [focused on new ways of constructing clothes].
- (Item 19) Which of the following is an illegal method of getting around quotas? [Transshipping through countries with quota availability]
- (Item 20) To curtail evasion, new country-of-origin laws in the United States have determined that a garment originates where [the first manufacturing steps are taken].
- (Item 21) Section 807 of the tariff schedule encourages foreign [contracting].
- (Item 22) Small retailers are less likely to take advantage of [foreign commissionaires and agents].
- (Item 23) Which of the following methods of penetrating foreign markets is used least? [Direct exports]
- (Item 29) [Trade deficit] is the amount by which the value of exports exceed imports.
5. To acquire a realistic understanding of the opportunities and responsibilities in the fashion industry.
- (Item 13) The part of the fashion industry that produces trims, threads, zippers, and interfacing [secondary suppliers].
- (Item 15) [Converter] is a firm that buys or handles the greige good from mills and contracts them out to finishing plants to have them finished.

CT 3204 - Textile Science

1. Define various textile terms and use them in proper context.

(No Related Subtest Items)

2. Distinguish between natural and man-made fibers, and among various members of each class.

- (Item 12) When protein fibers and cellulosic fibers are compared [protein fibers are weaker when wet].
- (Item 13) When natural and man-made fibers are compared [man-made] fibers are chemically modified by 15% weight through copolymerization to achieve specific properties for end use suitability.
- (Item 21) Which of the following fiber types could be used in a garment designed to meet the Federal Trade Commission flammability standards for children's sleepwear? [modacrylic]
- (Item 22) Which of the following fibers is the weakest fiber, especially when wet, but has good elongation and recovery, and high moisture regain? [Wool]
- (Item 23) Which of the following fibers is stiff, will not burn, and requires special handling? [Glass]
- (Item 24) Which of the following fibers is appropriate for apparel, household, and industrial uses and has the following properties: not acid resistant, not sunlight resistant, melts with burning? [Nylon]
- (Item 25) Which of the following fibers is lighter than water and does not absorb water, is resistant to acids and bases, but can be attacked by some organic cleaning solutions? [Olefin]
- (Item 26) Which man-made fiber has poor dimensional stability and is weaker when wet? [Rayon]

3. Relate properties of the various fibers to their chemical structures.

(No Related Subtest Items)

4. Evaluate various physical and chemical tests for fiber identification with consideration of inherent limitations.

(No Related Subtest Items)

5. Distinguish between spun (staple) and filament yarns, complex and miscellaneous yarns.

(No Related Subtest Items)

6. Use a yarn classification system to select a yarn type suitable for a textile product end use.

(No Related Subtest Items)

7. Distinguish among nonwoven, woven, and knit fabrics.

(Item 15) Raschel is a fabric using which of the following textile constructions? [Knit]

(Item 17) Which of the following descriptions corresponds with a needlepunch nonwoven construction? [Mechanical binding of fiber webs]

8. Relate the properties and principal uses of the various types of fabric constructions to consumer, industrial, and high performance applications.

(Item 14) Which of the following fabrics has a rib weave and can be used for outer-wear to resist wind and water penetration? [Poplin]

9. Recognize certain fabric finishes and coloration methods. Understand the chemical structures of finishes and dyes in relation to affinity for substrates and processing procedures.

(Item 18) Which of the following dye classes may require processing with carriers? [Disperse]

(Item 20) Which of the following dye classes has a chemical reaction with the textile fibers? [Reactive]

(Item 28) A process which smooths and straightens fabric when it is passes between heated rollers is [calendering].

10. Recognize the mechanical, aesthetic, functional effects of processing fibers, yarns, fabrics with selected finishes and coloration and the effects of these treatments on appearance and performance.

(Item 27) A fabric which has been treated with a chemical that increases the strength and luster in cotton and linen fabrics has been [mercerized].

(Item 29) A finishing process in which two layers of ribbed fabric are pressed together off grain is [moireing].

11. Analyze selected textile periodicals, according to specified guidelines, in relation to new developments in the textile industry.

(No Related Subtest Items)

12. Identify the names of common fabrics by physical characterization.

(Item 16) Which of the following fabrics has a complex or novelty weave with a combination of plain, twill, satin weaves, and a raised pattern? [Brocade]

(Item 19) Which of the following fabrics has a weft pile woven construction with pile stacked in vertical columns referred to as "wales"? [Corduroy]

**For CT 1004 and CT 3204 two subtest items were not related to any of the objectives.

APPENDIX D

Item Analysis for
CT 1004 Clothing and People

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
1	High Low	15 15	2 1	2 4	<u>11</u> <u>10</u>			21	.70	1	.067
2	High Low	15 15	<u>13</u> <u>4</u>	0 3	2 8			17	.57	9	.60
3	High Low	15 15	<u>15</u> <u>12</u>	0 2				27	.90	3	.20
4	High Low	15 15	<u>14</u> <u>7</u>	1 6				21	.70	7	.47
5	High Low	15 15	2 7	<u>13</u> <u>6</u>	0 2	0 0		19	.63	6	.40
6	High Low	15 15	0 1	<u>15</u> <u>9</u>	0 1	0 0	0 4	24	.80	6	.40
7	High Low	15 15	1 1	1 5	<u>13</u> <u>9</u>	0 0		22	.73	4	.27
8	High Low	15 15	1 7	<u>14</u> <u>4</u>	0 3	0 1		18	.60	10	.67
9	High Low	15 15	0 1	0 3	0 1	<u>14</u> <u>8</u>	1 2	22	.73	6	.40
10	High Low	15 15	0 8	<u>15</u> <u>6</u>				21	.70	9	.60
11	High Low	15 15	2 6	13 5	<u>0</u> <u>1</u>	0 3		1	.033	- 1	-.067
12	High Low	15 15	<u>12</u> <u>10</u>	1 2	0 1	0 1	2 1	22	.73	2	.13
13	High Low	15 15	1 4	0 2	0 1	<u>4</u> <u>0</u>	10 8	4	.13	4	.27
14	High Low	15 15	0 1	2 5	1 6	0 1	<u>12</u> <u>3</u>	15	.50	9	.60

Item Analysis for
CT 1004 Clothing and People (Continued)

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
15	High	15	1	0	0	1	<u>13</u>	23	.77	3	.20
	Low	15	1	1	1	2	<u>10</u>				
16	High	15	1	2	2	<u>8</u>		12	.40	4	.27
	Low	15	3	5	3	<u>4</u>					
17	High	15	0	0	0	<u>15</u>		27	.90	3	.20
	Low	15	0	1	2	<u>12</u>					
18	High	15	0	0	<u>15</u>			23	.77	7	.47
	Low	15	1	6	<u>8</u>						
19	High	15	<u>15</u>	0				26	.87	4	.27
	Low	15	<u>11</u>	4							
20	High	15	2	5	<u>7</u>	1		10	.33	4	.27
	Low	15	4	6	<u>3</u>	2					

Item Analysis for
CT 2404 Basic Apparel Assembly Lecture

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
1	High	11	<u>9</u>	0	1	1		18	.82	0	0.0
	Low	11	<u>9</u>	1	0	0					
2	High	11	0	0	<u>11</u>	0		20	.91	2	.18
	Low	11	0	1	<u>9</u>	1					
3	High	11	<u>11</u>	0	0	0		21	.96	1	.09
	Low	11	<u>10</u>	1	0	0					
4	High	11	0	4	1	<u>6</u>		10	.45	2	.18
	Low	11	1	6	0	<u>4</u>					
5	High	11	<u>7</u>	1	3	0		10	.45	4	.36
	Low	11	<u>3</u>	3	5	0					
6	High	11	1	<u>6</u>	3	1		6	.27	6	.55
	Low	11	3	<u>0</u>	7	1					
7	High	11	<u>9</u>	0	2	0		15	.68	3	.27
	Low	11	<u>6</u>	0	4	1					
8	High	11	0	0	1	<u>10</u>		12	.55	8	.73
	Low	11	2	1	6	<u>2</u>					
9	High	11	5	<u>6</u>	1	1		10	.45	2	.18
	Low	11	5	<u>4</u>	0	1					
10	High	11	1	4	<u>4</u>	2		6	.27	2	.18
	Low	11	3	1	<u>2</u>	5					
11	High	11	<u>5</u>	2	4	0		9	.41	1	.09
	Low	11	<u>4</u>	4	2	1					
12	High	11	<u>8</u>	2	1	0		13	.59	3	.27
	Low	11	<u>5</u>	4	0	2					
13	High	11	<u>10</u>	0	1	0		15	.68	5	.45
	Low	11	<u>5</u>	3	1	3					
14	High	11	0	1	2	<u>8</u>		15	.68	1	.09
	Low	11	1	2	1	<u>7</u>					

Item Analysis for
CT 2404 Basic Apparel Assembly Lecture (Continued)

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
15	High	11	3	<u>8</u>	0	0		10	.45	6	.55
	Low	11	4	<u>2</u>	3	2					
16	High	11	<u>6</u>	5	0	0		9	.41	3	.27
	Low	11	<u>3</u>	4	3	1					
17	High	11	0	1	<u>10</u>	0		15	.68	5	.45
	Low	11	1	5	<u>5</u>	0					
18	High	11	2	<u>7</u>	0	2		8	.36	6	.55
	Low	11	4	<u>1</u>	2	4					
19	High	11	<u>8</u>	0	0	3		11	.50	5	.45
	Low	11	<u>3</u>	5	1	2					
20	High	11	2	1	5	<u>3</u>		4	.18	2	.18
	Low	11	4	2	4	<u>1</u>					

Item Analysis for
CT 2414 Basic Apparel Assembly Laboratory

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
1	High	8	0	0	<u>8</u>	0		15	.94	1	.13
	Low	8	1	0	<u>7</u>	0					
2	High	8	<u>7</u>	0	0	1		12	.75	2	.25
	Low	8	<u>5</u>	0	0	3					
3	High	8	0	<u>6</u>	1	1		7	.44	5	.63
	Low	8	1	<u>1</u>	5	1					
4	High	8	0	0	0	<u>8</u>		13	.81	3	.38
	Low	8	0	2	1	<u>5</u>					
5	High	8	<u>7</u>	1	0			13	.81	1	.13
	Low	8	<u>6</u>	1	1						
6	High	8	0	0	<u>8</u>	0		13	.81	3	.38
	Low	8	0	2	<u>5</u>	1					
7	High	8	0	1	0	1	<u>6</u>	11	.69	1	.13
	Low	8	0	2	0	1	<u>5</u>				
8	High	8	0	<u>7</u>	0	1		13	.81	1	.13
	Low	8	0	<u>6</u>	0	2					
9	High	8	1	<u>7</u>	0	0		10	.63	4	.50
	Low	8	3	<u>3</u>	0	2					
10	High	8	0	0	0	<u>8</u>	0	16	1.00	0	0.0
	Low	8	0	0	0	<u>8</u>	0				
11	High	8	0	<u>6</u>	0	2		8	.50	4	.50
	Low	8	3	<u>2</u>	0	3					
12	High	8	0	0	<u>8</u>	0		14	.88	2	.25
	Low	8	0	0	<u>6</u>	2					
13	High	8	0	0	0	<u>8</u>		9	.56	7	.88
	Low	8	3	3	1	<u>1</u>					
14	High	8	<u>8</u>	0				14	.88	2	.25
	Low	8	<u>6</u>	2							
15	High	8	0	0	4	<u>4</u>		10	.63	-2	-.25
	Low	8	0	0	2	<u>6</u>					

Item Analysis for
CT 2414 Basic Apparel Assembly Laboratory (Continued)

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
16	High	8	<u>8</u>	0	0	0		15	1.00	0	0.0
	Low	8	<u>8</u>	0	0	0					
17	High	8	0	2	2	<u>4</u>		5	.31	3	.38
	Low	8	1	2	4	<u>1</u>					
18	High	8	0	0	<u>8</u>	0		14	.88	2	.25
	Low	8	0	1	<u>6</u>	1					
19	High	8	0	0	<u>4</u>	4		5	.31	3	.38
	Low	8	1	1	<u>1</u>	5					
20	High	8	<u>5</u>	1	2			7	.44	3	.38
	Low	8	<u>2</u>	3	3						

Item Analysis for
CT 2604 Introduction to the Fashion Industry

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
1	High	7	<u>7</u>	0	0	0		12	.86	2	.29
	Low	7	<u>5</u>	0	0	2					
2	High	7	0	<u>7</u>	0	0		13	.93	1	.14
	Low	7	1	<u>6</u>	0	0					
3	High	7	<u>7</u>	0	0	0		13	.93	1	.14
	Low	7	<u>6</u>	0	1	0					
4	High	7	1	<u>6</u>	0	0		10	.71	2	.29
	Low	7	2	<u>4</u>	1	0					
5	High	7	0	0	<u>2</u>	5		3	.21	1	.14
	Low	7	1	1	<u>1</u>	4					
6	High	7	<u>4</u>	2	0	1		6	.43	2	.29
	Low	7	<u>2</u>	3	0	2					
7	High	7	0	<u>7</u>	0	0		14	1.00	0	0.0
	Low	7	0	<u>7</u>	0	0					
8	High	7	3	0	0	<u>4</u>		7	.50	1	.14
	Low	7	4	0	0	<u>3</u>					
9	High	7	2	1	<u>2</u>	1		3	.21	1	.14
	Low	7	3	2	<u>1</u>	1					
10	High	7	0	0	1	<u>6</u>		8	.57	4	.57
	Low	7	3	2	0	<u>2</u>					
11	High	7	<u>7</u>	0	0	0		12	.86	2	.29
	Low	7	<u>5</u>	1	0	1					
12	High	7	<u>7</u>	0	0	0		9	.64	5	.71
	Low	7	<u>2</u>	4	0	1					
13	High	7	0	0	2	<u>5</u>		7	.50	3	.43
	Low	7	0	0	5	<u>2</u>					
14	High	7	1	<u>4</u>	1	1		5	.36	3	.43
	Low	7	0	<u>1</u>	3	3					
15	High	7	0	0	<u>3</u>	4		4	.29	2	.29
	Low	7	1	0	<u>1</u>	5					

Item Analysis for
CT 2604 Introduction to the Fashion Industry (Continued)

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
16	High	7	<u>7</u>	0	0	0		14	1.00	0	0.0
	Low	7	<u>7</u>	0	0	0					
17	High	7	2	<u>5</u>	0	0		8	.57	2	.29
	Low	7	1	<u>3</u>	1	2					
18	High	7	<u>7</u>	0	0	0		13	.93	1	.14
	Low	7	<u>6</u>	0	1	0					
19	High	7	0	<u>7</u>	0	0		14	1.00	0	0.0
	Low	7	0	<u>7</u>	0	0					
20	High	7	0	<u>7</u>	0	0		13	.93	1	.14
	Low	7	1	<u>6</u>	0	0					

Item Analysis for
CT 3204 Textile Science

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
1	High	3	<u>3</u>	0				6	1.00	0	0.0
	Low	3	<u>3</u>	0							
2	High	3	2	<u>1</u>				1	.17	1	.33
	Low	3	3	<u>0</u>							
3	High	3	<u>2</u>	0	3	0		3	.50	1	.33
	Low	3	<u>1</u>	1	1	0					
4	High	3	0	<u>3</u>				6	1.00	0	0.0
	Low	3	0	<u>3</u>							
5	High	3	<u>2</u>	0	1	0	0	3	.50	1	.33
	Low	3	<u>1</u>	0	0	1	1				
6	High	3	0	<u>3</u>	0	0	0	4	.67	2	.67
	Low	3	1	<u>1</u>	1	0	0				
7	High	3	0	1	<u>2</u>	0	0	4	.67	0	0.0
	Low	3	0	1	<u>2</u>	0	0				
8	High	3	0	<u>3</u>	0			4	.67	2	.67
	Low	3	1	<u>1</u>	1						
9	High	3	1	1	0	1	0	1	.17	1	.33
	Low	3	<u>0</u>	2	1	0	0				
10	High	3	0	0	0	0	<u>3</u>	5	.83	1	.33
	Low	3	0	0	0	1	<u>2</u>				
11	High	3	<u>3</u>	0	0	0	0	5	.83	1	.33
	Low	3	<u>2</u>	0	0	1	0				
12	High	3	1	0	0	<u>2</u>		3	.50	1	.33
	Low	3	1	1	0	<u>1</u>					
13	High	3	<u>2</u>	0	0	0	1	2	.33	2	.67
	Low	3	<u>0</u>	1	0	1	1				
14	High	3	0	0	0	0	<u>3</u>	6	1.00	0	0.0
	Low	3	0	0	0	0	<u>3</u>				

Item Analysis for
CT 3204 Textile Science (Continued)

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
15	High	3	0	<u>1</u>	<u>2</u>	0	0	4	.50	1	.33
	Low	3	0	1	<u>1</u>	1	0				
16	High	3	0	<u>3</u>	0	0	0	4	.67	2	.67
	Low	3	1	<u>1</u>	0	1	0				
17	High	3	0	0	<u>3</u>	0	0	4	.67	2	.67
	Low	3	0	1	<u>1</u>	1	0				
18	High	3	0	<u>2</u>	0	0	1	3	.50	1	.33
	Low	3	1	<u>1</u>	0	0	1				
19	High	3	0	0	0	<u>3</u>	0	4	.67	2	.67
	Low	3	0	0	1	<u>1</u>	1				
20	High	3	0	0	1	0	<u>2</u>	3	.50	1	.33
	Low	3	0	0	2	0	<u>1</u>				

APPENDIX E

Item Analysis for
Core Knowledge Exam - Consolidated

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
1	High Low	7 7	0 1	1 1	<u>6</u> <u>5</u>			11	.79	1	.14
2	High Low	7 7	<u>6</u> <u>6</u>	0 0	1 1			12	.86	0	0.0
3	High Low	7 7	<u>7</u> <u>7</u>	0 0				14	1.00	0	0.0
4	High Low	7 7	<u>6</u> <u>4</u>	1 3				10	.71	2	.29
5	High Low	7 7	0 1	<u>7</u> <u>6</u>	0 0	0 0		13	.93	1	.14
6	High Low	7 7	0 0	<u>6</u> <u>5</u>	0 0	0 0	1 2	11	.79	1	.14
7	High Low	7 7	0 0	1 2	<u>6</u> <u>5</u>	0 0		11	.79	1	.14
8	High Low	7 7	0 0	<u>7</u> <u>7</u>	0 0	0 0		14	1.00	0	0.0
9	High Low	7 7	0 0	0 3	0 0	<u>7</u> <u>4</u>	0 0	11	.79	3	.43
10	High Low	7 7	1 3	<u>6</u> <u>4</u>				10	.71	2	.29
11	High Low	7 7	0 1	7 4	<u>0</u> <u>2</u>	0 0		2	.14	-2	- .29
12	High Low	7 7	<u>7</u> <u>7</u>	0 0	0 0	0 0	0 0	14	1.00	0	0.0
13	High Low	7 7	0 3	0 0	0 0	<u>2</u> <u>1</u>	5 3	3	.21	1	.14
14	High Low	7 7	0 0	0 0	2 4	0 0	<u>5</u> <u>3</u>	8	.57	2	.29

Item Analysis for
Core Knowledge Exam - Consolidated (Continued)

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
15	High Low	7 7	1 0	0 0	0 0	0 1	<u>6</u> <u>6</u>	12	.86	0	0.0
16	High Low	7 7	0 0	2 6	4 1	<u>1</u> <u>0</u>		1	.07	1	.14
17	High Low	7 7	0 0	0 1	0 1	<u>7</u> <u>5</u>		12	.86	2	.29
18	High Low	7 7	0 1	0 3	<u>7</u> <u>3</u>			10	.71	4	.57
19	High Low	7 7	<u>7</u> <u>7</u>	0 0				14	1.00	0	0.0
20	High Low	7 7	0 0	0 2	<u>7</u> <u>4</u>	0 1		11	.79	3	.43
21	High Low	7 7	<u>6</u> <u>5</u>	0 0	0 0	1 2		11	.79	1	.14
22	High Low	7 7	0 3	0 0	<u>7</u> <u>3</u>	0 1		10	.71	4	.57
23	High Low	7 7	<u>7</u> <u>7</u>	0 0	0 0	0 0		14	1.00	0	0.0
24	High Low	7 7	0 1	0 0	0 1	<u>7</u> <u>5</u>		12	.86	2	.29
25	High Low	7 7	<u>5</u> <u>2</u>	0 3	2 2	0 0		7	.50	3	.43
26	High Low	7 7	0 0	<u>3</u> <u>0</u>	4 6	0 1		3	.21	3	.43
27	High Low	7 7	<u>6</u> <u>7</u>	0 0	1 0	0 0		11	.79	-1	- .14
28	High Low	7 7	0 1	0 0	0 4	<u>7</u> <u>2</u>		9	.64	5	.71

Item Analysis for
Core Knowledge Exam - Consolidated (Continued)

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
29	High Low	7 7	4 3	<u>0</u> <u>1</u>	0 1	3 2		1	.07	-1	- .14
30	High Low	7 7	1 1	0 2	<u>5</u> <u>3</u>	1 1		8	.57	2	.29
31	High Low	7 7	<u>6</u> <u>4</u>	0 0	1 2	0 1		10	.71	2	.29
32	High Low	7 7	<u>6</u> <u>5</u>	1 1	0 0	0 1		11	.79	1	.14
33	High Low	7 7	<u>7</u> <u>6</u>	0 0	0 0	0 1		13	.93	1	.14
34	High Low	7 7	0 1	0 2	2 2	<u>5</u> <u>2</u>		7	.50	3	.43
35	High Low	7 7	0 0	<u>6</u> <u>7</u>	0 0	1 0		13	.93	-1	- .14
36	High Low	7 7	<u>6</u> <u>5</u>	1 2	0 0	0 0		11	.79	1	.14
37	High Low	7 7	0 1	0 0	<u>7</u> <u>6</u>	0 0		13	.93	1	.14
38	High Low	7 7	0 0	<u>7</u> <u>7</u>	0 0	0 0		14	1.00	0	0.0
39	High Low	7 7	<u>7</u> <u>6</u>	0 0	0 0	0 1		13	.93	1	.14
40	High Low	7 7	0 1	0 1	1 4	<u>6</u> <u>1</u>		7	.50	5	.71
41	High Low	7 7	0 1	0 0	<u>7</u> <u>5</u>	0 1		12	.86	2	.29
42	High Low	7 7	<u>5</u> <u>2</u>	0 0	0 0	2 5		7	.50	3	.43

Item Analysis for
Core Knowledge Exam - Consolidated (Continued)

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
43	High Low	7 7	0 0	<u>3</u> <u>4</u>	2 3	2 0		7	.50	-1	- .14
44	High Low	7 7	0 0	0 2	0 0	<u>7</u> <u>5</u>		12	.86	2	.29
45	High Low	7 7	<u>6</u> <u>6</u>	0 0	1 0			12	.86	0	0.0
46	High Low	7 7	1 2	0 0	<u>6</u> <u>3</u>	0 2		9	.64	3	.43
47	High Low	7 7	0 0	1 2	0 0	0 1	<u>6</u> <u>4</u>	10	.71	2	.89
48	High Low	7 7	0 1	<u>7</u> <u>6</u>	0 0	0 0		13	.93	1	.14
49	High Low	7 7	0 1	<u>7</u> <u>5</u>	0 0	0 1		12	.86	2	.29
50	High Low	7 7	0 0	0 1	0 0	<u>7</u> <u>6</u>	0 0	13	.93	1	.14
51	High Low	7 7	0 1	<u>7</u> <u>5</u>	0 0	0 1		12	.86	2	.29
52	High Low	7 7	0 0	0 0	<u>7</u> <u>6</u>	0 1		13	.93	1	.14
53	High Low	7 7	0 2	0 2	2 0	<u>5</u> <u>3</u>		8	.57	2	.29
54	High Low	7 7	<u>7</u> <u>6</u>	0 1				13	.93	1	.14
55	High Low	7 7	0 2	0 0	4 2	<u>3</u> <u>3</u>		6	.43	0	0.0
56	High Low	7 7	<u>7</u> <u>7</u>	0 0	0 0	0 0		14	1.00	0	0.0

Item Analysis for
Core Knowledge Exam - Consolidated (Continued)

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
57	High	7	1	1	0	<u>5</u>		9	.64	1	.14
	Low	7	1	1	1	<u>4</u>					
58	High	7	0	0	<u>7</u>	0		12	.86	2	.29
	Low	7	0	2	<u>5</u>	0					
59	High	7	1	0	<u>3</u>	3		4	.29	2	.29
	Low	7	3	0	<u>1</u>	3					
60	High	7	<u>5</u>	0	2			7	.50	3	.43
	Low	7	<u>2</u>	0	5						
61	High	7	<u>7</u>	0	0	0		11	.79	3	.43
	Low	7	<u>4</u>	0	0	3					
62	High	7	1	<u>6</u>	0	0		9	.64	3	.43
	Low	7	4	<u>3</u>	0	0					
63	High	7	<u>7</u>	0	0	0		14	1.00	0	0.0
	Low	7	<u>7</u>	0	0	0					
64	High	7	1	<u>6</u>	0	0		13	.93	-1	- .14
	Low	7	0	<u>7</u>	0	0					
65	High	7	1	1	<u>4</u>	1		5	.36	3	.73
	Low	7	0	5	<u>1</u>	1					
66	High	7	<u>2</u>	2	0	3		3	.21	1	.14
	Low	7	<u>1</u>	3	0	3					
67	High	7	0	<u>7</u>	0	0		14	1.00	0	0.0
	Low	7	0	<u>7</u>	0	0					
68	High	7	5	0	0	<u>2</u>		7	.50	3	.73
	Low	7	2	0	0	<u>5</u>					
69	High	7	0	1	<u>3</u>	3		4	.29	2	.29
	Low	7	1	5	<u>1</u>	0					
70	High	7	0	1	0	<u>6</u>		10	.71	2	.29
	Low	7	2	1	0	<u>4</u>					

Item Analysis for
Core Knowledge Exam - Consolidated (Continued)

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
71	High	7	<u>4</u>	3	0	0		7	.50	1	.14
	Low	7	<u>3</u>	2	0	2					
72	High	7	<u>6</u>	1	0	0		10	.71	2	.29
	Low	7	<u>4</u>	2	1	0					
73	High	7	1	0	4	<u>2</u>		4	.29	0	0.0
	Low	7	2	0	3	<u>2</u>					
74	High	7	0	<u>5</u>	2	0		8	.57	2	.29
	Low	7	0	<u>3</u>	4	0					
75	High	7	0	1	<u>5</u>	1		9	.64	1	.14
	Low	7	0	0	<u>4</u>	3					
76	High	7	<u>7</u>	0	0	0		12	.86	2	.29
	Low	7	<u>5</u>	0	0	2					
77	High	7	1	<u>5</u>	0	1		6	.43	4	.57
	Low	7	2	<u>1</u>	0	4					
78	High	7	<u>7</u>	0	0	0		14	1.00	0	0.0
	Low	7	<u>7</u>	0	0	0					
79	High	7	0	<u>7</u>	0	0		14	1.00	0	0.0
	Low	7	0	<u>7</u>	0	0					
80	High	7	0	<u>7</u>	0	0		13	.93	1	.14
	Low	7	1	<u>6</u>	0	0					
81	High	7	<u>7</u>	0				13	.93	1	.14
	Low	7	<u>6</u>	1							
82	High	7	5	<u>2</u>				2	.14	2	.29
	Low	7	7	<u>0</u>							
83	High	7	<u>3</u>	0	1	3		6	.43	0	0.0
	Low	7	<u>3</u>	0	3	1					
84	High	7	0	<u>7</u>				13	.93	1	.14
	Low	7	1	<u>6</u>							

Item Analysis for
Core Knowledge Exam - Consolidated (Continued)

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
85	High	7	<u>6</u>	0	0	0	1	9	.64	3	.43
	Low	7	<u>3</u>	0	1	2	1				
86	High	7	0	<u>5</u>	1	1	0	9	.64	2	.29
	Low	7	0	<u>4</u>	2	1	0				
87	High	7	0	0	<u>7</u>	0	0	10	.71	4	.57
	Low	7	0	0	<u>3</u>	1	3				
88	High	7	1	<u>6</u>	0			11	.79	1	.14
	Low	7	1	<u>5</u>	1						
89	High	7	<u>5</u>	1	1	0	0	8	.57	2	.29
	Low	7	<u>3</u>	3	1	0	0				
90	High	7	0	0	1	0	<u>6</u>	10	.71	2	.29
	Low	7	0	0	1	2	<u>4</u>				
91	High	7	<u>6</u>	0	1	0	0	10	.71	2	.29
	Low	7	<u>4</u>	0	2	1	0				
92	High	7	0	0	0	<u>7</u>		12	.86	2	.29
	Low	7	2	0	0	<u>5</u>					
93	High	7	<u>7</u>	0	0	0	0	10	.71	4	.57
	Low	7	<u>3</u>	2	0	0	2				
94	High	7	0	0	0	2	<u>5</u>	8	.57	2	.29
	Low	7	1	2	1	0	<u>3</u>				
95	High	7	1	1	<u>5</u>	0	0	8	.57	2	.29
	Low	7	0	2	<u>3</u>	2	0				
96	High	7	0	<u>7</u>	0	0	0	9	.64	5	.71
	Low	7	3	<u>2</u>	0	2	0				
97	High	7	0	0	<u>6</u>	1	0	10	.71	2	.29
	Low	7	0	1	<u>4</u>	2	0				
98	High	7	0	<u>6</u>	1	0	0	9	.64	3	.43
	Low	7	1	<u>3</u>	0	1	2				

Item Analysis for
Core Knowledge Exam - Consolidated (Continued)

Item No.	Group	Group Size	Response Alternatives					Total Right	Difficulty Index	(H-L)	Discrimin. Index
			a	b	c	d	e				
99	High	7	0	0	0	<u>7</u>	0	9	.64	4	.57
	Low	7	0	2	3	<u>2</u>	0				
100	High	7	0	0	0	0	<u>7</u>	7	.50	7	1.00
	Low	7	2	1	4	0	<u>0</u>				

APPENDIX F

Verification of the Key
for the
Clothing & Textiles Core Knowledge Exam

Exam Item Numbers	Source
10 - 12	Sproles, G.B. (1979). <u>Fashion: Consumer behavior toward dress.</u> Minneapolis: Burgess Publishing Company, pp. 26, 60-61.
13,15,91	Kefgen, M. & Touchie-Specht, P. (1986). <u>Individuality in clothing selection and appearance.</u> New York: Macmillan Publishing Company, pp. 71, 136, 532.
14	Phillips, J.W. & Staley, H.K. (1975). Sumptuary legislation in four centuries. In L.M. Gurel & M.S. Beeson (Eds.), <u>Dimensions of dress and adornment: A book of readings,</u> (pp. 33-36). Dubuque, Iowa: Kendall/Hunt Publishing Company, p. 33.
16,19,27	Horn, M.J. & Gurel, L.M. (1981). <u>The second skin.</u> Boston: Houghton Mifflin Co., pp. 3-5, 36-39, 160.
17	Packard, S., Winters, A.A., & Axelrod, N. (1983). <u>Fashion buying and merchandising.</u> New York: Fairchild Publications, p. 7.
18	Gurel, L.M. (1975). Anthropological perspectives of dress and adornment: Introduction. In L.M. Gurel & M.S. Beeson (Eds.), <u>Dimensions of dress and adornment: A book of readings,</u> (pp. 9-10). Dubuque, Iowa: Kendall/Hunt Publishing Company, p. 9.
20,29,73,75	Frings, G.S. (1987). <u>Fashion: From concept to consumer.</u> Englewood Cliffs: Prentice-Hall, Inc., pp. 60, 68, 79, 83.

Exam Item Numbers	Source
21,70-72,76-89	Jarnow, J. J. Guerreiro, M., & Judelle, B. (1987). <u>Inside the fashion business</u> . New York: Macmillan Publishing Co. pp. 12, 29, 37, 44, 134, 150, 155, 305, 309, 336, 339, 342, 343, 347, 354, 433, 433, 447.
22-24, 36, 96	Wingate, I.B. & Mohler, J.F. (1984). <u>Textile fabrics and their selection</u> . Englewood Cliffs: Prentice-Hall, Inc., pp. 8-10, 111, 177-178.
25, 74	Stone, E. & Samples, J.A. (1985). <u>Fashion merchandising: An introduction</u> . New York: McGraw-Hill Book Company, pp. 4, 49.
26	Veblen, T. (1975). Dress as an expression of the pecuniary culture. In L.M. Gurel & M.S. Beeson (Eds.), <u>Dimensions of dress and adornment: A book of readings</u> , (pp. 111-116). Dubuque, Iowa: Kendall/Hunt Publishing Company, pp. 111-116.
28	Horn, M.J. (1968). <u>The second skin</u> . Boston: Houghton Mifflin Company, p. 2.
30, 31	Verified by department faculty not teaching the course.
32, 33, 35	Glock, R.E. & Kunz, G.I. (1990). <u>Apparel manufacturing: Sewn product analysis</u> . New York: Macmillan Publishing Company, pp. 38, 110-116, 258.
34	Hollen, N., Saddler, J., Langford, A.L., & Kadolph, S.J. (1988). <u>Textiles</u> . New York: Macmillan Publishing Company, p. 258.
37	Tate, S.L. (1984). <u>Inside fashion design</u> . New York: Harper & Row Publishers, p. 204-205.

Exam Item Numbers	Source
38-39, 43, 92, 98, 100-101, 104	Joseph, M.L. (1981). <u>Introductory textile science</u> . New York: Holt, Rinehart, & Winston, Inc., pp. 50, 97, 127, 312, 314, 349, 353, 382.
40, 42, 44, 47, 49-51, 53-54, 56-58, 60-67	<u>Reader's digest complete guide to sewing</u> . (1978). Pleasantville, New York: The Reader's Digest Association, Inc., pp. 27, 28, 61-69, 72, 74, 144, 145, 147-150, 159, 202, 254, 256, 291, 317.
41	Stamper, A.A., Sharp, S.H., & Donnell, L.B. (1986). <u>Evaluating apparel quality</u> . New York: Fairchild Publications, p. 35.
45, 52, 68-69	Perry, P. (Ed.). (1970). <u>The vogue sewing book</u> . New York: Vogue Patterns, pp. 77, 105, 246-247.
46, 59, 55	Bane, A. (1973). <u>Creative clothing construction</u> . New York: McGraw-Hill Book Company, pp. 87, 403, 515.
48	Kallal, M.J. (1985). <u>Clothing construction</u> . New York: Macmillan Publishing Company, p. 247.
90	Mauldin, W.L. & Beeson, M.S. (1975). Reading the labels on apparel and household textiles. In L.M. Gurel & M.S. Beeson (Eds.), <u>Dimensions of dress and adornment: A book of readings</u> , (pp. 129-132). Dubuque, Iowa: Kendall/Hunt Publishing Company, p. 129.
93	Lyle, D.S. (1976). <u>Modern textiles</u> . New York: John Wiley & Sons, Inc., p. 39.
94-95, 97, 99, 102-103, 105-109	Pizzuto, J.J. (1980). <u>Fabric science</u> . New York: Fairchild Publications, pp. 32, 40, 43, 45, 171, 184, 237, 239, 242, 344, 346.

VITA

Valerie Littlejohn Giddings, the daughter of Mrs. Ethel G. Littlejohn and the late Horace C. Littlejohn, Sr. was born August 14, 1958 in Spartanburg, South Carolina. She graduated from Spartanburg High School in 1976. In May, 1980 she received a Bachelor of Science degree in Home Economics with a concentration in Clothing and Textiles from Bennett College in Greensboro, North Carolina.

In 1982 she married Ronald Giddings and in the same year received a Master of Science degree in Clothing and Textiles from Virginia Tech in Blacksburg, Virginia. In the Spring of 1987, she entered the doctoral program at Virginia Tech in clothing and textiles.

Her professional experiences are as follows:

1980 - 1982 Graduate Assistant in the Department of Clothing and Textiles, VPI&SU, Blacksburg, VA.

1982 - 1985 Instructor, Fashion Merchandising & Apparel Design, Department of Human Ecology, Hampton University, Hampton, VA.

1985 - 1986 Instructor, Fashion Merchandising & Apparel Design, Department of Home Economics, North Carolina Central University, Durham, NC.

1987 - Lecturer and Extension Specialist, Department of Clothing
Present and Textiles, Virginia Tech, Blacksburg, VA.

She is a member of Kappa Omicron Nu, Phi Upsilon Omicron, Alpha Kappa Mu, International Textile and Apparel Association, and Delta Sigma Theta Sorority, Incorporated.