

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

**AGREEMENT AMONG COLLEGE READING INSTRUMENTS AND
THEIR RELATION TO DEVELOPMENTAL COURSE PLACEMENT**

by

Mary Draga Jurkiewicz Shelor

A Thesis Submitted to the Faculty of the
DEPARTMENT OF LANGUAGE, READING AND CULTURE

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF EDUCATION

In the Graduate College

THE UNIVERSITY OF ARIZONA

1998

UMI Number: 9901696

UMI Microform 9901696
Copyright 1998, by UMI Company. All rights reserved.

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

THE UNIVERSITY OF ARIZONA ©
GRADUATE COLLEGE

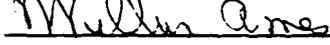
As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Marv Draga Jurkiewicz Shelor entitled Agreement Among College Reading Instruments and their Relation to Developmental Course Placement.

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Education



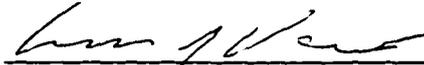
John Bradley

June 29, 1998
Date



Wilbur Ames

June 29, 1998
Date



William Valmont

6/29/98
Date

Date

Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.


Dissertation Director John Bradley

June 29, 1998
Date

STATEMENT OF AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Mary K. Sheeran

ACKNOWLEDGMENTS

I wish to express my gratitude to Dr. John Bradley, Chair of my Dissertation Committee, for all his help, patience and continued friendship during this long process. Thanks also to Dr. Bill Ames and Dr. Bill Valmont for serving on my committee and offering their time, suggestions, and encouragement.

Over the years, a number of people in Language, Reading and Culture have shared their knowledge, friendship and life experiences, and I wish to say thank you to Dr. Adela Allen, Dr. Arminda Fuentesvilla, Dr. Judy Mitchell and Dr. Richard Ruiz.

In addition, thank yous are extended to Pat Griesel, Ken Ritterspach, Jeanne Struck, Penny Turrentine, and many others who shared research, read drafts, discussed my study, or just checked in periodically offering a word of encouragement. Their support was invaluable.

My deepest love and appreciation go to my first teachers, my Mom and Dad, Ann and Bill Jurkiewicz, who taught me the value of education and hard work. More importantly, they taught me about the important things in life, such as God, family and friends, and blessed me with three sisters and brothers, who continue to challenge me and enrich my life.

Finally, I thank my son, Matthew Swanson Shelor, who has made this long journey with me. He was always there for me, encouraging me to continue and always believing in me, even when I didn't. I dedicate this work to him.

TABLE OF CONTENTS

LIST OF TABLES	7
ABSTRACT	9
CHAPTER 1 INTRODUCTION	11
Background of the Study	11
Statement of the Problem	20
Procedures for Collecting Data	20
Hypotheses	21
Justification for the Study	22
Definition of Terms	23
Assumptions	25
Limitations	26
CHAPTER 2 REVIEW OF THE LITERATURE	27
Introduction	27
Profile of the Typical College Student of the 90s	27
How Institutions of Higher Education Identify Underprepared Students	33
Reading Assessment Instruments	36
The Nelson-Denny Reading Test	38
The Degrees of Reading Power	43
The Turrentine/Bradley Literacy Testing Battery	46
A Comprehensive Assessment/Placement Plan	48
CHAPTER 3 PROCEDURES	53
The Southwestern Community College	53
The Community College 1--The Study Site	56
The Sample	57
The Developmental Reading Curriculum	59
Procedures for Gathering Data	62
The Reading Assessment Instruments	67
Analysis of the Data	73
CHAPTER 4 FINDINGS	76
Introduction	76
Description of the Analysis Procedure	76
Explanation of the Kappa	78
Correlational Findings	78
Extent of Agreement Findings	82
ANOVA Findings	93

TABLE OF CONTENTS--*Continued*

Qualitative Data	101
Ann	102
Guillermo	105
Liliana	108
 CHAPTER 5 SUMMARY, CONCLUSION, IMPLICATIONS AND RECOMMENDATIONS	 113
Summary	113
Conclusions	118
Implications	120
Recommendations	122
 APPENDIX A PARTICIPANT'S CONSENT FORM	 124
 APPENDIX B DREAMS READING PASSAGE AND TEST QUESTIONS	 126
 APPENDIX C PUBLISHER'S PERMISSION LETTER	 137
 REFERENCES	 139

LIST OF TABLES

Table		
3.1	Characteristics of the Sample Compared to SWCC and CC1 in Fall 1997	59
3.2	Criterion NDRT Raw Scores for Placement into Four SWCC Reading Courses	63
3.3	Criterion DRP Raw Scores for Placement into Four SWCC Reading Courses	65
3.4	Criterion TBLTB Raw Scores for Placement into four SWCC Reading Courses	66
3.5	Inter-Rater Reliability on the TBLTB-SA Format for 10 Subjects	68
3.6	Grade Equivalency for NDRT-Form E	69
4.1	Pearson Product-Moment Correlation Coefficients Between the Scores of the Tests	79
4.2	The Number of Students Correctly Placed, Underplaced and Overplaced by the DRP When Compared with the NDRT	83
4.3	The Number of Students Correctly Placed, Underplaced and Overplaced by the NDRT When Compared with the TBLTB-MC	85
4.4	The Number of Students Correctly Placed, Underplaced and Overplaced by the NDRT When Compared with the TBLTB-SA	86
4.5	The Number of Students Correctly Placed, Underplaced and Overplaced by the DRP When Compared with the TBLTB-MC	88
4.6	The Number of Students Correctly Placed, Underplaced and Overplaced by the DRP When Compared with the TBLTB-SA	89
4.7	The Number of Students Correctly Placed, Underplaced and Overplaced by the TBLTB-MC When Compared with the TBLTB-SA	90
4.8	Kappa Coefficients Between the Tests	91

LIST OF TABLES--*Continued*

4.9	Mean Scores and Standard Deviations Computed According to Placement by the NDRT When Computed with the DRP	94
4.10	Mean Scores and Standard Deviations of the TBLTB-MC and TBLTB-SA, According to Placement by the NDRT	94
4.11	Mean Scores and Standard Deviations of the TBLTB-MC and TBLTB-SA, According to Placement by the DRP	95
4.12	Analysis of Variance among the DRP Mean Scores for Students Placed into Three Developmental Reading Classes by the NDRT	98
4.13	Analysis of Variance among the TBLTB-MC Mean Scores for Students Placed into Three Developmental Reading Classes by the NDRT	98
4.14	Analysis of Variance among the TBLTB-SA Mean Scores for Students Placed into Three Developmental Reading Classes by the NDRT	99
4.15	Analysis of Variance among the TBLTB-MC Mean Scores for Students Placed into Three Developmental Reading Classes by the DRP	100
4.16	Analysis of Variance among the TBLTB-SA Mean Scores for Students Placed into Three Developmental Reading Classes by the DRP	100

ABSTRACT

The purpose of this study was to investigate the assessment instruments that a southwestern community college (SWCC) used to evaluate its incoming students' academic literacy skills and to determine how well the results of those tests placed students in either college-level or developmental reading courses. The three college-level academic literacy instruments that were investigated were: the Nelson-Denny Reading Test (NDRT), the Degrees of Reading Power (DRP), and the Turrentine/Bradley Literacy Testing Battery (TBLTB), consisting of two response types: the TBLTB Multiple-Choice (TBLTB-MC) and the TBLTB Short-Answer (TBLTB-SA). Interrelationships were computed for reliability, concurrent validity and extent of agreement for placement into the three developmental reading courses.

The three tests were administered to 88 students placed into the three developmental reading courses, Reading 071, Reading 081 and Reading 091, taught by the researcher during the Fall 1997 semester. The NDRT was used by the college for placement purposes. The DRP and the TBLTB were given during the first and second weeks of the semester for comparison purposes.

Pearson Product-Moment correlations indicated that there were low to moderately low positive correlations among all the tests at the .01 level of confidence.

The Kappas obtained between pairings of the three tests demonstrated that two of the tests showed concurrent placement validity: (1) the NDRT with the TBLTB-MC and

(2) the TBLTB-MC with the TBLTB-SA. The extent of agreement exceeded the .01 level of confidence. The other tests did not show sufficient extent of agreement for reading developmental course placement purposes.

The ANOVAS demonstrated significant mean differences among students placed into the three SWCC developmental reading classes by: (1) the NDRT and the DRP, (2) the NDRT and the TBLTB-MC and (3) the NDRT and the TBLTB-SA at the .01 level of confidence. Two other combinations of tests showed significance at the .05 level of confidence: (1) the DRP and the TBLTB-MC and (2) the DRP and the TBLTB-SA.

Although the three tests did not have extent of agreement sufficient for placement into the three developmental reading classes, all three tests did agree that this sample of students did not possess college-level reading abilities.

CHAPTER 1

INTRODUCTION

The purpose of this chapter is to present the (1) background of the study, (2) statement of the problem, (3) procedures for collecting data, (4) hypotheses, (5) justification for the study, (6) definition of terms, (7) assumptions, and (8) limitations.

Background of the Study

There are a variety of factors that contribute to the need for developmental reading programs in community colleges. One major factor is related to the open-door admission's policy that community colleges typically have. Community colleges do not usually require a high school diploma or other criteria (such as a particular minimum GPA or a minimum score on the SAT or ACT) for entrance. Anyone, 18 years of age or older, who wants to take classes may attend a community college. Such students could be first time students, transfer students, students with disabilities, students concurrently enrolled in high school or returning students who have been out of school for many years in hopes of acquiring new skills toward improved employment opportunities. This open-door admission policy allows all members of a community to attend college without regard for their previous academic and/or literacy experiences and preparation. Thus, this admission policy permits individuals with limited literacy and academic skills to become community college students. As a consequence, many of those students are not prepared

for the rigors of college courses and may require additional support in terms of developmental courses to assist them to succeed.

The typical community college student in the 1950s was a recent high school graduate, Anglo, male, and attending school full time. Now, a typical community college student is increasingly older, a female and an ethnic or racial minority who often is working and raising a family while going to school part time. In addition, many community college students are foreign born, so they may not have acquired their elementary or secondary education in the United States and are unfamiliar with the American educational system. More importantly, many may not have English as their primary language which is the language of instruction in community colleges. Thus, most community colleges have high student diversity which includes linguistic, cultural, economic and academic differences. Such diversity poses a challenge to institutions of higher learning to assist their students in being successful.

Wide diversity in academic preparation, and in particular, literacy competency poses a real challenge to our community colleges. Key tasks facing community college reading faculty include the accurate assessment of students' literacy strengths and skills and the correct placement of students into developmental literacy courses or regular college courses. This is no easy task in light of the highly diverse student population of most community colleges.

Research suggests that accurate placement in developmental courses is very important to student success because a mastery of basic academic literacy ability lays the

groundwork for successful performance in college courses (Levitz & Noel, 1995; Morante, 1989; The College Board, 1983). Early success in reading courses, beginning with accurate and effective placement measures, can influence student performance in other key areas of the college curriculum, including writing, foreign languages, the humanities, the sciences, social sciences and applied mathematics, all of which have become increasingly reading-based (Ediger, 1991; Hills, Hirsch & Subhiyah, 1990). If a literacy assessment instrument is valid, then it will identify those students who are ready for college-level courses as well as those who need some type of intervention (e.g., a developmental reading course) to be successful in college-level course work.

Research also has revealed that the provision of early and appropriate literacy interventions by community colleges can lower student drop out rates, a major problem in this country (Fullilove & Treisman, 1990; Roueche, 1984; Roueche & Snow, 1973). The most recent comprehensive investigation of drop-out rates and college retention rates for colleges and universities was done by EXXON and the National Center for Developmental Education (Boylan, Bonham, & Bliss, 1992). It found that 60% of all entering freshmen drop out of universities and 72% to 76% of all entering freshmen drop out of community colleges before completing their courses of study. Although many factors, such as lack of financial resources, family problems, work related conflicts, other responsibilities and low self-esteem, have been documented to account for this startling drop out statistic, one consistent overriding factor has been that many failing students are underprepared for the academic demands of college (Astin, 1975; Pascarella & Terenzini,

1991; Tinto, 1987, 1988). That is, drop outs almost always lack the basic skills needed to read and comprehend college-level texts. As a consequence, such students are more likely to drop out because they are not sufficiently prepared academically for college-level course work.

Even though the need for developmental reading courses has been evident since institutions of higher learning were first established in this country in the seventeenth century, the need continues to grow. According to the Report of the National Center for Education Statistics (NCES, 1990), the number of students coming to college academically underprepared grew 114% from 1970 to 1989 (Chumchal, 1996). The impact of the growth of academically underprepared students is demonstrated by the fact that today there are some 2500 developmental education programs operating in the more than 3000 colleges and universities across the country (Boylan, Bingham, & Cockman, 1992).

However, the EXXON/National Association of Developmental Education Study (Boylan, Bonham, & Bliss, 1992) found that even though students who are initially underprepared have lower success rates than students judged to be better academically prepared, they are less likely to fail if they participate in developmental literacy programs. In addition, students who participate in developmental literacy programs obtain higher grades than students who do not participate in them, and they have higher rates of retention, almost as high as those of the initially better-prepared students. Community

college developmental literacy courses often give many initially underprepared students the boost they need to succeed.

The need for developmental literacy courses and the success of these courses places paramount importance on choosing the most valid and reliable instrument to assess the academic literacy skills of incoming students. If students are assessed accurately and subsequently are placed appropriately in academic and literacy support programs, then they have a greater chance of staying in school and ultimately graduating (Hills, Hirsch, & Subhiyah, 1990; Morante, 1989).

To be valid, college reading assessment instruments should evaluate students' academic literacy and study skills that parallel such tasks used in lower-division college-level courses. If it is determined that students can read well enough to comprehend the textbooks that are used in first year college-level courses, then they are deemed ready to be placed in college-level course work. If the assessment reveals that students are not prepared to successfully read such textbooks, then appropriate developmental courses should be provided. A major problem encountered in this process is that some researchers and educators question the validity of most of the currently used reading assessment instruments in terms of adequately assessing the actual reading demands of college-level course work (Farr & Carey, 1986; Flippo & Caverly, 1991; Mason, 1994).

Since most community colleges have no set academic criteria for entrance, over 50% see a need to assess their incoming students' basic academic and literacy skills (Boylan, Bonham, & Bliss, 1994). To accomplish this, 94% of those colleges use

norm-referenced standardized tests to evaluate their incoming college students (El-Khawas, 1991). Standardized tests are typically constructed by experts who carefully check tests for validity and reliability. Norm-referenced tests are generally administered to large numbers of individuals who represent cross-sections of the intended populations for each test. Norms or normative sample scores are provided so an individual's performance can be measured and compared in relation to that of a peer group population. This comparison is valid only if the characteristics of the students being tested parallel those of the test's normative sample.

The advantages of using standardized norm-referenced tests are numerous. They are easy to administer and can be scored quickly and objectively which makes them time efficient. They are also useful for gaining indications of the overall performance of students while providing a peer reference group for comparison purposes.

Unfortunately, there are also important drawbacks connected with the use of norm-referenced tests. The norm-referenced tests that most community colleges typically use often have been normed on populations of students very different from today's community college students. In addition, norm-referenced tests are typically limited in what they can assess. For example, most norm-referenced tests assess reading vocabulary knowledge in isolation and out of context and assess reading comprehension with relatively short reading passages. Critics wonder if these norm-referenced test formats are comparable to actual college-level reading activities since college students usually need to determine vocabulary meaning through context and comprehend lengthy texts of

many pages (Sternberg, 1991). The main criticism of norm-referenced reading tests is that they measure only a limited range of literacy skills, providing only a gross measure of reading ability (Maxwell, 1994). Because of this, they cannot identify the specific strengths and weaknesses of students in terms of college-level literacy competency.

Criterion-referenced tests are also sometimes used by community colleges to assess entry-level college students. Criterion-referenced tests provide student test performance information relative to a specific standard or predetermined criterion regarding success or failure with the particular domain of knowledge or abilities. They typically are designed to identify specific strengths and weaknesses of students and can be designed to assess mastery or non-mastery of specific academic literacy skills.

There are numerous advantages to using criterion-referenced tests. Since criterion-referenced tests can be geared to specific objectives and even student characteristics, they can be used to make formative evaluations, such as diagnosing specific strengths and weaknesses of individual students and determining their participation in selected programs. Normally, they are graded locally, so immediate feedback with control over the use and interpretation of the results are possible. Overall, they can yield more specific and clearer information than norm-referenced tests.

However, criterion-referenced tests have their disadvantages as well. Finding a published academic literacy test that assesses the domain of a particular reading course or predicts to a particular community college curriculum is problematic. If a reading department decides to develop its own criterion-referenced reading test, then the task can

be time consuming and costly. For example, during the various phases of test development, considerable leadership, expertise and coordination are necessary. Writing unambiguous, unbiased and discriminating test items is a complex process. And constructing a criterion-referenced test that is both reliable and valid requires a level of sophisticated psychometrics that many faculty do not possess (Morante, 1987; Prus & Johnson, 1994).

A more cost-effective alternative might be an informal, criterion-referenced inventory, where a college text is used as a source of reading material for the test. Students are given a particular passage from a college text and then are asked a series of comprehension questions, both multiple-choice and short-answer questions, to evaluate their literacy skills. Students are asked to identify main ideas and details, identify the type of text structure, or compare and contrast issues discussed in the text. These higher level thinking activities match more appropriately with the competencies in critical reading and thinking that educators suggest college students should be able to do in college-level courses.

Using a college text as the basis for constructing a criterion-referenced literacy placement test improves its content validity, since students are taking a test constructed of materials that are sampled from the types of texts that they are asked to read for their college-level courses. Research has found that using regular textbooks for assessment purposes is more motivating and less demeaning to students (Maxwell, 1994).

There are other advantages to the use of informal, criterion-referenced tests. Since any subject area college textbook can be used, the tests can be molded to fit the specific needs of any student. In addition, the reading comprehension questions can be given either orally or in written form. Including some type of short-answer comprehension questions in such a test also allows students to construct knowledge rather than just identify right answers as they do with multiple-choice items. In addition, short-answer items provide information regarding a student's writing competency as well as his/her reading competency. Students are often required to perform higher level thinking activities when they write their answers to comprehension questions.

An over-riding disadvantage of this type of test is that it can be very time-consuming both in administration and evaluation. However, these tests provide a great deal of information about students' specific academic related literacy skills and strengths, enhancing the probability of providing the type of information appropriate for making informed decisions about a student's academic placement in either college-level courses or developmental courses.

In summary, (1) greater numbers of students are coming to community colleges lacking the academic literacy skills needed for success; (2) these students need to be identified so that appropriate literacy interventions can be provided; and (3) the college reading assessment instruments currently being used are problematic regarding their validity as predictors of academic literacy competence.

Statement of the Problem

The purpose of this study was to investigate the assessment instruments that a southwestern community college (SWCC) used to evaluate its incoming students' academic literacy skills and to determine how well the results of these tests placed students in either college-level or developmental courses. Three college-level academic literacy instruments were investigated: the Nelson-Denny Reading Test (NDRT), the Degrees of Reading Power (DRP) and the Turrentine/Bradley Literacy Testing Battery (TBLTB). This study investigated the interrelationships of the NDRT, the DRP and the TBLTB for reliability, concurrent validity and extent of agreement for placement into developmental reading courses. Specifically, the purposes for this research study were to determine:

1. The correlation among the NDRT, the DRP and TBLTB.
2. The extent of agreement among the NDRT, the DRP and TBLTB in terms of placement into SWCC's developmental reading courses.

Procedures for Collecting Data

A sample of students (n=88) taking classes at a community college during the Fall semester of 1997 participated in the study. These students were also taking a developmental reading course taught by the researcher. All three reading assessment instruments, the NDRT, the DRP and the TBLTB were administered as a pre-test to all of the students. In addition, a survey gathering data on participants' age, sex, ethnicity,

place of birth, high school, first language and enrollment status was administered during the first three weeks of classes.

Hypotheses

1. There is no correlation among all three tests, the NDRT, the DRP and the TBLTB.
 - A. There is no relationship between the scores of the NDRT and the DRP.
 - B. There is no relationship between the scores of the NDRT and the TBLTB.
 - C. There is no relationship between the scores of the DRP and the TBLTB.
 2. There is no extent of agreement among all of the tests on placement.
 - A. There is no agreement between the NDRT and the DRP regarding course placement.
 - B. There is no agreement between the NDRT and the TBLTB regarding course placement.
 - C. There is no agreement between the DRP and the TBLTB regarding course placement.
 - D. There are no increases among the mean scores of the TBLTB from the lowest developmental reading course to the highest course when the subjects are placed in the three courses by the NDRT.
 - E. There are no increases among the mean scores of the TBLTB from the lowest developmental reading course to the highest course when the subjects are placed in the three courses by the DRP.
-

Justification for the Study

Most community college faculties are aware of the need to assess students' incoming literacy skills. In fact, assessing students' basic academic skills has become a key strategy in improving student success (Bray, 1987). However, the literacy information that community colleges have regarding incoming students is often inadequate. It cannot be assumed that just because a student has graduated from high school, he or she is adequately prepared for college-level courses. There is little continuity from one high school to another on the contents of English courses and how they were measured. In addition, many incoming students may have been away from high school for a number of years, and they no longer remember the skills and knowledge that they studied in high school.

Theoretically, an assessment program utilizes valid and reliable instruments to assess incoming students' literacy skills. But in reality, this is not necessarily true. Community colleges should periodically reevaluate currently used literacy assessment instruments to determine if they do in fact, assess skills adequately, so appropriate guidance in placing students in either college-level or developmental courses is assured. Research confirms that improved assessment and placement procedures ultimately help to reduce high drop-out rates in college (Fullilove & Treisman, 1990; Roueche, 1984; Roueche & Snow, 1973).

Another factor that makes this study important is that community colleges expend a large portion of their yearly budget on providing developmental reading courses for

underprepared students. Although no firm estimates of the costs of community college developmental education seem to exist, a range of \$2-10 billion in costs across the country has been estimated (Abraham, 1987). Given the amount, one would expect better results.

In summary, there appears to be a need to investigate and develop appropriate literacy assessment instruments that can identify academic literacy skills of incoming community college students and place them appropriately. If students have adequate academic literacy skills then they can be placed in college-level course work. If they need an intervention program, then the type of instruction needed should be determined from the results of the assessment instrument.

Definition of Terms

The following terms are defined to give clarity to the use of them in this study.

College Academic Reading Skills. Skills needed for the reading of college materials including strategies related to decoding, reading comprehension, speed, and critical thinking skills that are required to understand and evaluate information contained in college-level textbooks.

Criterion-referenced Tests. Tests that compare a person's performance with a standard or mastery of a predetermined criterion for a particular domain such as reading or writing.

Concurrent Validity. The extent of relationship between two separate measures or tests of the same domain such as reading and/or writing.

Developmental. Generally defined as the process of growth, unfolding, or activation.

Thus “normal” growth is developmental. In the context of this study, it refers to the academic and/or literacy knowledge and skills needed to succeed in college-level courses.

Developmental Reading Courses. Reading instruction designed to improve students’ competencies in the basic areas of literacy such as vocabulary, comprehension and critical thinking.

Informal Tests. Tests that employ a wide variety of procedures to assess reading performance, such as paper and pencil tests, oral tests, observations, etc.

Internal Reliability. Internal reliability is achieved when each item correlates highly to the total test score.

Literacy. The ability to read and write in a language. For this study, the reading and writing ability sufficient for success in college-level course work.

Norm-referenced Tests. Tests that compare the performance of an individual to that of a reference or normative group defined in terms of age, grade, etc. Normative groups may be local, regional or national.

Norms. Statistics that define the performance of specific groups, such as age or grade.

Parallel Form Reliability. The relationship between scores in two or more forms of a test with each form consisting of a set of unique items designed to measure the same domain such as reading, writing, etc.

Placement. The assignment of a person to an appropriate course or academic program in accordance with his/her aims, capabilities, readiness, educational background, and/or aspirations.

Predictive Validity. A type of measurement validity that deals with the question of whether an assessment forecasts a person's performance on a future task. Specifically, in this study, it is the extent to which a college-level literacy test predicts future performance in a particular college course, set of courses or overall success in college in terms of GPA

Reliability. Reliability may be determined in terms of correlation between a testing and retesting of the same form of a test, correlates across two or more forms of the same test and/or through an analysis of the consistency of the items within a form of the test.

Validity. The extent to which a test measures what it purports to measure. In this study, the extent to which a test measures the ability to read and/or write at a college level.

Assumptions

The following assumptions were necessary to conduct this study:

1. It is assumed that the time of day that the tests were administered had no significant effect on the performance of students.
 2. It is assumed that the academic literacy assessments used in this study are valid for use with second language learners.
-

Limitations

SWCC, the community college that was the focus of this study is situated in the southwest, close to the Mexican border. Because of its location, there is a high percentage of Hispanic students who typically enroll in SWCC. The study is limited in that it may not be generalizable to other community college campuses across the country with different populations of students.

There are many reading assessment instruments available for use by community colleges to use with incoming first year students. It would be difficult to investigate all of these instruments. Only two of the most widely used ones were included in this study. These two instruments were studied because they were already being used by SWCC.

Another limitation of this study is the choice of the reading passage for the Turrentine/Bradley Literacy Testing Battery (TBLTB). The TBLTB used a chapter from a college psychology text book. Students may have performed very differently on a chapter from another college text in psychology or a text covering a different area such as history, music or science. Also, if a student had a considerable amount of prior knowledge in the tested content or had a high interest in that topic, then he/she might score higher on that instrument. Using a variety of content test passages with the same group of students would undoubtedly improve the validity of this study, but it would also make the study unmanageable in size, so it was decided to use only this one passage.

CHAPTER 2

REVIEW OF THE LITERATURE

Introduction

The following issues were noted in Chapter 1 as problematic when considering the assessment and placement of incoming community college students: (1) a greater number of students are coming to community colleges lacking the academic literacy skills needed for success, (2) these students who are underprepared need to be identified so that interventions can be provided, and (3) the college reading assessment instruments currently being used are problematic.

Topics included in this chapter are: (1) the profile of the typical college student, (2) how institutions of higher education identify underprepared students, (3) reading assessment instruments, (4) the Nelson-Denny Reading Test, (5) the Degrees of Reading Power, (6) the Turrentine/Bradley Literacy Testing Battery, and (7) comprehensive assessment/placement plans.

Profile of the Typical College Student of the 90s

The typical college student 50 years ago was male (60%) and white (97%) (Lucas, 1996). The students who attended college were ranked in the upper third of their high school graduating classes, ranging in age from 18 to 22 years of age. The majority of these students were single and attended school full time. They lived on campus either in a dormitory, fraternity house or sorority house rather than at home or in apartments off

campus. The liberal arts degree was the typical program of choice and most students completed their studies in four years.

The typical college student of today can be quite different from the typical college student of 50 years ago. For example, the majority of college students are female (55%) and the students who attend college today are the most racially diverse generation of students ever in the history of the U.S. (Kerlin & Britz, 1994). Students are also older with the majority of them in their mid-twenties or late-twenties. For instance, 44% are over the age of 25 years of age, and 11% are 40 years or over (U.S. Department of Education, 1996).

Since many college students are older, married and raising families, school is often not the central focus of their lives. Other responsibilities such as work take precedence. Fifty-four percent of students work and 24% of them work full time while attending college. Almost half (43%) of today's students attend college only part time (U.S. Department of Education, 1996b).

When students were asked about their educational and career goals, the majority of them said they were in search of job security and aspired toward financial independence (Levine & Cureton, 1998). They stated that they were going to college to acquire marketable skills and to prepare for a profession in business, education or a health field. Many of the students interviewed reported that they had seen parents and friends laid off from big companies due to downsizing, file bankruptcy or become homeless during their lifetime. They were interested in a career that would offer them security in

today's volatile economy. This is in significant contrast to the students of the 50s, 60s and 70s who were studying for a Liberal Arts Degree.

A major obstacle that may hinder some current college students from attaining their goal of completing college may be attributed to the fact that they are underprepared for college (Roueche & Baker, 1987). Maxwell (1979) defined underprepared learners as students "whose skills, knowledge and academic abilities are significantly below those of the 'typical' students in the college curriculum in which they are enrolled" (p. 2).

Roueche and Roueche (1993) expanded the definition by categorizing these students as being at risk because they are not only underprepared, but they also work 30 or more hours per week, have little family support, are first-generation college attenders and have "failure expectations." Failure to possess the basic academic literacy skills necessary to read and comprehend college-level texts is only a part of the factor of underpreparedness.

Since the early eighties, a variety of studies (Evangelauf, 1985; Kirst, 1997; Lederman, Ryzewic & Ribaud, 1985; Morante, Faskow & Menditto, 1984; National Center for Education Statistics, 1995) have been done to document the number of students coming to college underprepared. Most of these studies have reported that from 29% to 39% of entering students do not possess basic proficiencies in reading, writing and mathematics needed to perform successfully at the college level (Abraham, 1987; Grable, 1988; Mason, 1994).

Morante, Faskow, and Menditto (1984) have done extensive research in the state of New Jersey assessing the academic skills of incoming college students and studying

the effectiveness of the New Jersey College Basic Skills Placement Test (NJCBS) that is used to evaluate them. The NJCBS is a common statewide instrument administered to all incoming college students. The results of this test have indicated that thousands of incoming students lack basic skills proficiencies. Using the NJCBS, it was found that 71% of recent high school graduates demonstrated deficiencies in verbal skills, 62% in math and 86% in elementary algebra.

Lederman, Ryzewic, and Ribaud (1985) surveyed faculty in over 1200 colleges and universities throughout the United States about their perceptions of incoming students and whether or not they thought these students were academically prepared for college. The survey results revealed that 85% of four-year faculty members perceived incoming freshmen as underprepared for college course work and 90% of two-year faculty members had the same perceptions of their students. Results of basic skill assessments of incoming students supported the instructors' perceptions. It was found that 28% of the students needed some sort of assistance in reading, 31% in writing and 32% in mathematics. Astin (1991) completed a similar study six years later and found the same results. Of the 35,000 faculty surveyed in over 400 U.S. colleges and universities, they reported that the surveyed faculty perceived only 27% of incoming students were academically well-prepared for college.

A study done by Evangelauf (1985) during the academic year 1983-1984 in the state of California found that 60% of incoming students failed to demonstrate entry-level reading and writing competencies and that 53% of the students could not perform basic

math. Ten years later, Krist's (1995) study noted the same problem. About half of the incoming freshmen into the California State University (CSU) System needed some form of remedial work in both English and math.

Tennessee (1991) and Texas (Skinner & Carter, 1987) conducted similar statewide studies of incoming college students in an attempt to ascertain the number of students needing developmental assistance. In Tennessee, a combination of several instruments were used to evaluate incoming students' academic abilities and the obtained results showed that 50% of the students needed remedial math, 37% needed writing assistance, and 38% needed reading assistance. In Texas, incoming students were evaluated at all public state colleges and universities. Of the 110,000 students assessed, only 30,000 or approximately 27% of the students could read and write well enough to perform successfully in college-level courses. When only community colleges students were tested, it was found that 70% of the students did not have the necessary skills to be successful in college, with 60% of the students demonstrating the need for assistance in reading, 57% in writing and 46% in math.

Over the last 15 years, the National Center for Education Statistics (NCES) in the U.S. Department of Education has compiled data on college-level remedial education. Data were collected during the 1983-1984, the 1989-1990, and the 1994-1995 academic school years. During the 1983-1984 academic school year, it was reported that 25% of all incoming freshmen were enrolled in developmental English, reading and/or math courses. In addition, 82% of all colleges and universities reported the need to offer remedial

courses. Six years later, 30% of all college freshmen were enrolled in at least one remedial course, with 58% of the students taking a reading course (Roueche & Roueche, 1991). During that academic school year, 74% of all colleges and universities reported having to offer remedial courses for underprepared students, an 8% decrease from the previous study.

In the Fall of 1995, NCES (Ignash, 1997) surveyed over 3000 two- and four-year colleges in the U.S. and found that 100% of public two-year and 80% of public four-year colleges and universities offered remedial programs. In the private sector, 63% of two- and four-year colleges offered remedial courses. Colleges with open-door admission policies and high minority enrollments showed even higher percentages of remedial programs. As many researchers have noted, it is difficult to determine if trends are increasing or decreasing. But it is reasonable to conclude that the results from these national studies and other studies show that students underprepared for college-level course work seem to permeate higher education.

In contrast, other researchers, (Maxwell, 1997; Roueche & Roueche, 1993; Van, 1992) have observed that the underpreparedness of incoming freshmen is not a new phenomenon. To substantiate their conclusions, these researchers have chronicled how the American educational system has experienced influxes of underprepared students during different periods throughout the history of the United States. Each time as higher education became available to another sector of society, i.e., women, the working class, Blacks, etc., colleges and universities had to accommodate new and different populations

of students who were not ready for college-level work. For instance, during the “Jacksonian Period” from 1824-1848 (Maxwell, 1994) higher education became a reality for the “common man.” There was a shortage of tradesmen, engineers, and agricultural specialists throughout the country, so new colleges were established in practically every state by means of the passage of the Morrill Act of 1862. Since few men had attended school past grammar school, and secondary schools were just in their early beginnings, (Wyatt, 1992) this new population of students came unprepared for the educational demands of college in the early 19th century.

Other historical events that brought influxes of unprepared and underprepared students into colleges and universities were: (1) the passage of the GI Bill after World War II, giving governmental support to veterans to attend college; (2) the passage of the Civil Rights Act of 1964 and the Higher Education Act of 1965, offering increased admissions to the educationally disadvantaged, minorities and women (Van, 1992); and (3) the passage of the American Disabilities Act of 1990. Boylan would call this “cycles of influx by non-traditional students” as different components of American society were offered the possibility of higher education (Payne & Lyman, 1992). Whether we perceive underpreparedness as an old issue or a new one, it is a significant problem that colleges and universities have at this time.

How Institutions of Higher Education Identify Underprepared Students

Colleges and universities typically identify underprepared students through some form of assessment. The assessment instrument used and the criteria for determining the

level of underpreparedness varies greatly from institution to institution. However, there appears to be agreement that students lacking basic academic skills are coming to college and some kind of assessment is necessary to identify their abilities to handle the rigors of college course work.

A study done by the American College Testing Program and the American Association of Community and Junior Colleges (AACJC) found that more than 75% of both urban and non-urban institutions of higher education required some form of basic skills testing or submission of scores on a norm-referenced standardized test for all entering students (1985). In 1988, the State Higher Education Executive Officers adopted a position paper that advocated among other things, that all institutions of higher education assess all incoming students for the purpose of determining their incoming basic academic skills (Cress, 1996). By 1989, the process of assessing incoming students' basic skills had been institutionalized by 65% of all colleges and universities in the U.S. and another 19% had plans to implement such measures (Krotseng & Pike, 1991). Almost all community colleges required incoming students to be assessed in the basic skills area (El-Khawas, 1989).

Through assessment, students' basic academic skills and deficiencies are identified, which provides the basis for advising students regarding the courses they might take throughout their college programs. This information is vital to students since research has found that assessing and placing students in college courses that best match their academic preparation has become a key strategy for improving and increasing

student success (Bray, 1987; McNabb, 1990). Retaining students in college is the ultimate goal. Yet, of all students who enter college, fewer than half graduate from their first institution of enrollment (Boylan & Bonham, 1992; Tinto, 1987).

One might question that since the majority of incoming students are high school graduates it would be reasonable to assume that most of them should be adequately prepared for the academic demands of college. However, over the years, it has become apparent that high school grades and the possession of a high school diploma may be unreliable indicators of academic preparation for college, because of grade inflation, lack of consistency in academic standards, inaccurately administered tests, and other factors. As the state of New Jersey learned, students who had completed high school and earned a diploma were not necessarily ready for college (Morante et al., 1989). In fact, many of the students requiring remediation obtained above average scores on the SAT.

Another sector of the community college population that might require basic skills assessment so appropriate advisement can be made consists of students who did not finish high school, students with disabilities, or older students who may have been out of school for many years. All of these students may need special services in addition to assessment to assist them in being successful college students.

Another area where assessment has shown an incredible growth has been within the universities and colleges themselves for accountability purposes (Cress, 1996). With increasing pressure from external and internal constituencies, such as accreditation bodies, legislators, taxpayers and parents, the need to prove growth, efficiency and

effectiveness has quadrupled. Thus, assessing students' basic skills provides not only an avenue for sorting and placing students into courses, but a way to establish performance standards and goals, and answer to the public as the stakeholders demand results (McMillan, 1994).

Reading Assessment Instruments

Reading is one of the three basic skills (reading, writing and mathematics) that is needed to be successful in college (Mason, 1994). Twenty years ago, it was found that a student's reading level was highly correlated with performance and/or survival in 93% of college subject areas (MacDouglass, 1977). Thus, it is assumed that being a highly efficient and strategic reader is required for successful progress in a lower-division college curriculum (Hills, Hirsch, & Subhiyah, 1990). Yet, it is estimated that one-third to one-half of all incoming freshmen and reentering students need some form of reading assistance so they can comprehend college-level texts (Kerstiens, 1990; McWhorter, 1994; Ross & Roe, 1986).

Published standardized tests constitute the majority of the reading tests that have been used to assess college students' reading comprehension and were used to determine students' lack of adequate reading skills. As early as the 1930s, standardized tests were criticized for possessing low reliability and validity measures among other things. Yet, they continue to be used today. In addition, it has been contended that standardized reading tests provide little information to make accurate judgments about students' abilities and do not adequately reflect actual reading demands of academic course work in

college. Ideally, college reading tests should assess readers' abilities to deal with the academic demands of college-level course work (Smith & Jackson, 1985).

There seems to be consensus that the current standardized reading tests used to assess incoming students' literacy skills are inadequate for that purpose (Malinowski, 1988; Seybert, 1991). Among the reasons why they are inadequate are that: (1) they focus more on the literal and factual information that can be lifted from the text without reading than on critical thinking (Kiah, 1993); (2) they may be inappropriate for today's diverse student population (National Commission on Testing and Public Policy, 1990), and (3) they cannot test the process of reading, only the products of reading comprehension (Wood, 1989).

Standardized norm-referenced and criterion-referenced tests are designed to measure the product of reading, instead of the process that one engages in when reading (Farr & Carey, 1986; Johnston, 1983). Current reading tests look upon reading as a compilation of isolated skills, and it is assumed that if a student accurately reads the exact words of the text, comprehension will be accomplished. It is contended that reading involves much more than the simple act of word recognition. A reader brings his/her own meanings to the text and as the reader reads, there is an interaction between the text of the writer and the schema of the reader. They create meaning together, and that meaning is influenced by the reader's culture, background knowledge and experiences. This new theory of reading, constructivism, cannot be tested by the existing reading tests, particularly those that are standardized (Wood, 1989). That is why there is a need for

assessment measures that focus on how readers construct meaning. Standardized tests usually consist of short passages followed by multiple-choice questions. These passages are nothing like the reading assignments that college students need to perform in college. College students are required to read long, difficult textbooks written in various formats which include titles, authors, captions, graphs, references, etc. These text factors are rarely included in standardized test reading passages. In addition, the multiple-choice format used by almost all standardized tests requires students to only pick the right answer, rather than create an answer appropriate in terms of their culture and prior knowledge. Adding a writing component to a reading test would allow students to compare and contrast issues or synthesize information demonstrating higher level thinking skills while they are reading.

The Nelson Denny Reading Test

The Nelson-Denny Reading Test (NDRT) (1981) is widely used to assess incoming students' reading abilities. The NDRT is a test that assesses high school and college students' academic achievement in reading comprehension, vocabulary development and reading rate. It is a standardized, norm-referenced, timed test, that is easy to give and comparatively inexpensive to use. The authors suggest that the test can be used to identify not only students who may need special help in reading, but also identify superior students who could profit from placement in advanced classes (Brown, Bennet & Hanna, 1981). However, Tierney (1985a) has questioned the NDRT's ability to

be used to make such fine interpretations as identifying both superior and underprepared students.

Flippo (1980c) and Maxwell (1991) agreed with Tierney. Flippo categorized the NDRT as a survey test that is inappropriate for diagnosing students' individual strengths and weaknesses in reading. Survey tests are sometimes referred to as screening instruments providing information about students' general levels of reading ability. If a survey test is used to identify a student who does have a reading problem, then a diagnostic test would also have to be administered to determine the student's specific reading problems.

Maxwell (1991) added that the NDRT is a norm-referenced test, which shows how a student scored in comparison with a particular normative peer group. She stated that the NDRT may have value at the college level when used for screening purposes, but care should be exercised in interpreting its scores for purposes of individual diagnosis. She concluded that what is needed is a reading placement test that is criterion-referenced, so that specific student reading strengths and weaknesses can be identified.

Studies conducted by Cummins (1981) and Van Blerkom and Van Blerkom (1981) supported the notion that interpreting the scores on the NDRT for diagnostic purposes was problematic. Cummins (1981) investigated the effects of altered directions for taking the NDRT. Students were tested on Form E following the standard procedures of the NDRT and then immediately tested on Form F. Although before taking Form F, students were instructed to go straight to the questions and just locate the answers in the

passage, but not read the passage. The percent of correct answers increased significantly from Test E to Test F, with no intervention, except the added instructions preceding Form F, which can be described as a test-taking strategy. This study showed that a single test-taking technique can alter test scores on a multiple-choice test, putting into question the reliability of the test.

Van Blerkom and Van Blerkom (1981) did a similar study after noting improved scores in posttest scores. They gave instructions to students that they were to increase their reading speed on the posttest by just filling in the bubbles as quickly as possible, near the end of the testing time. Posttest score improvement was explained by the fact that more test items were completed on the second test. However, the percentage of attempted items that were correct decreased slightly (not significantly) from the pretest to the posttest. This is another study putting into question the reliability of the NDRT.

Another point of concern has been that the NDRT is a strictly timed test. Speed is important in aptitude tests that measure capability. These types of tests are timed so that predictions for success in college can be made about the better and brighter students (Morante, 1989). Yet, when timed tests are used to assess basic skills, the fact that a test is timed decreases its validity because the discriminate ability of the test is reduced. Kerstiens (1986) found little justification for time constraints on reading tests and described the issue in more practical ways. Kerstiens observed that timed tests bear little relation to the kinds of reading freshmen students do in college. Normally, there is ample time to read and reread as a student tries to comprehend a passage in a text. So, timed

reading tests are not realistic measures of reading as it occurs in college and are poor indicators of what students can read and comprehend when they have time to think and question and apply meaning. In addition, Deem (1980) has speculated that timed reading tests may be biased against culturally and linguistically different students, who may have a whole different meaning of the concept of time than that held by the creators of the test.

Not only is the NDRT difficult, but its time limits are unrealistic. Mason (1994) recommended that a solution to this problem could be to do one of two things: (1) lengthen the time limit on the test or do away with the time limits altogether, or (2) do not alter the time allocation, but train students in the most vulnerable groups (those who have been identified through the research to be affected by the time factor) in strategic skills which will rapidly accelerate their test-taking rate.

Farr (1972) pointed out that a large percentage of students not having the time to complete the test is a sign that the test lacks discriminant validity which is a serious deficiency in any reading test. The NDRT lacks discriminant validity in the sense that on the average, college freshmen correctly complete only 33 of the 100 vocabulary items and that the norming sample is above the college freshman level. This means that the NDRT is unable to diagnose specific reading skill strengths and weaknesses, especially at the low end of the norming sample (Raygor, 1978).

One of the most comprehensive studies of the NDRT was a six-year longitudinal study (Wood & Bennett, 1990) with more than 11,000 freshmen in a Midwestern, state-supported university from 1983 to 1988. The purpose of the study was to determine the

relationships between the NDRT-E scores and other indicators of academic success such as college grade point average, high school GPA, high school rank in class and the American College Testing (ACT) or Scholastic Aptitude Test (SAT) scores. The researchers found that the NDRT-E correlated positively with the ACT composite score ($r=.69$) and the SAT Verbal score ($r=.77$). A lower correlation was found with high school GPA ($r=.26$) and college GPA ($r=.22$). The results of this study confirmed the idea that the NDRT provided little predictive value for academic success in college in terms of individual scores on the vocabulary and comprehension subtests.

The NDRT may be too difficult for most community college students. Content validity of the NDRT was achieved by choosing vocabulary words and reading passages that ranged in difficulty from grades 9 through 16. Researchers (Brown, Bennett, & Hanna, 1981) noted that the vocabulary and reading passages did exhibit an appropriate range of difficulty levels. Other researchers (Kerstiens, 1986; Tierney, 1985) reported that some of the passages were archaic and unlike high school textbooks. Other researchers concurred (Boylan, 1983; Perkins, 1984) who said that the passages were too hard. Some suggested that the NDRT should be used only with selected college students because the instrument is a good tool for identifying superior students (Raygor, 1978; Rounds, 1984; Van Blerkom & Van Blerkom, 1981). Raygor (1978) emphatically stated poorly prepared students should not be tested with the NDRT because it could be a defeating and frustrating experience.

In summary, the NDRT may be used for screening purposes, but care should be exercised in interpreting scores for purposes of individual diagnosis. It appears that it might be appropriate to reexam the use of the NDRT as an entry-level assessment instrument (Perkins, 1984).

The Degrees of Reading Power

The Degrees of Reading Power (DRP) (1983) was developed by the New York State Department of Education, Touchstone Applied Science Associates (TASA) and the College Board, and was designed to measure a student's ability to read and understand English prose at different levels of difficulty. The DRP does not measure subskills, but instead "measures a single, holistic objective in reading--the ability to read with comprehension." (The College Board, 1983).

The DRP is untimed and consists of eleven reading passages written in a combination cloze/multiple-choice format. Each passage of approximately 350 words, taken from *Encyclopedia Britannica*, has seven omitted words at regular intervals. For each omitted word there are five multiple-choice response options to choose among. All the information necessary to respond to each cloze/multiple-choice item is provided in the passage. The difficulty of the items is linked to the difficulty of the passage. All possible responses are "semantically plausible and syntactically correct" so students must use context and understand the surrounding prose to choose the correct word among the five provided (Enright, 1987).

The *DRP Handbook* (Touchstone Applied Science Associates, 1995) noted the following characteristics of the test:

1. Students must be able to read and understand the DRP test items, which are carefully designed, so they can answer the questions correctly.
2. Students are not required to have any prior knowledge to answer the questions. All information is contained in the reading passage.
3. All responses are common words, even in the more difficult reading passages.
4. The test is designed to deter guessing.
5. Item difficulty is linked to test difficulty.

These characteristics have encouraged reviewers (Bruning, 1985; Kibby, 1981; Stoker, 1987) to review the test positively, judging it to be among the best-conceived and carefully constructed measures of reading comprehension available. Strickland (Wood, 1989) like others was excited about the modified cloze format, since it resembles real reading. Overall, the test's reliability and validity has been rated to be generally good and the test is considered to use excellent testing methodology.

Bruning (1985) stated that the DRP shows good reliability, convergent-discriminate validity and construct validity. The DRP can predict performance with reasonable accuracy and is sensitive to gain. Criticisms focused on the absence of a technical manual (now available), poor test administration instructions, and lack of information on the norming population. Since this critique, the DRP has been normed again in 1994.

The 1994 norming sample included some 35,000 students in grades 1-12 in all schools in the State of New York. Bias was examined with respect to gender, ethnicity and socioeconomic status and no effects were found. Looking at the ethnic breakdown of students just in grade 12, that might be comparable to the participants in this study, it was found that: 65.4% were White, 6.6% Spanish, 1.8% Asian, .2% Native American, 12.3% Black and 12.8% not given. The ethnic breakdown corresponded to the national ethnic breakdown. However, Ysseldyke (1985) criticized the DRP because the norming population did not have adequate Black and Hispanic representation.

ESL students were also included in the DRP norming sample to see if its passages contained cultural content that might bias the test results. Since the test is designed so all content information is in the text, students do not have to rely on memory or experience to choose the correct response word. Thus, it is probable that the DRP is free of cultural bias and is a valid measure for ESL students.

Hanna (1985) pointed out positive characteristics of the test which included passage independence, norm referencing and reliability. However, he cautioned colleges about the inability of the DRP to place students with appropriate instructional materials, as did Carver (1985). Carver was more forceful in his conclusions and stated that the DRP is not valid on the basis that it cannot do what it purports to do. He said the DRP appeared to have large internal inconsistencies that render it invalid for matching reading materials to students varying in reading ability. Yet Carver did concede that the DRP did

evaluate how individuals compared with each other in reading achievement and indicated if remedial assistance was needed.

Turrentine/Bradley Literacy Testing Battery

The Turrentine/Bradley Literacy Testing Battery (TBLTB) (1997), an informal criterion-referenced reading instrument, was created to assess the literacy abilities of incoming first-year university students (Turrentine & Bradley, 1997). The goal of the TBLTB is to provide an early and accurate assessment of the literacy strengths and weaknesses of students so appropriate interventions could be provided if needed. The purpose of the study was to determine if the TBLTB was a reliable and valid assessment instrument that could be used for assessing undergraduate college students' academic and literacy abilities.

There are two forms of the TBLTB each consisting of a long reading passage from an undergraduate psychology text. The TBLTB has one passage on Dreams (Form 1) and another passage on Cognition (Form 2). Each reading passage is about six pages long. Both passages are followed by a set of thirteen multiple-choice questions and ten short-answer questions. One hundred and thirty-eight students were tested and Pearson Product-Moment Correlation Coefficients were calculated between the two forms (Dreams and Cognition) and across response types (multiple choice and short answer). The extent of agreement between the two forms (Dreams and Cognition) and across response types (multiple-choice and short-answer) was also determined and a qualitative analysis of three cases was done.

It was found that the correlation coefficients between Forms 1 and 2 of the TBLTB across response types were moderate. However, the extent of agreement across forms and response types was relatively high indicating that the tests were reliable assessment instruments for the purpose of placement. The tests also proved to be valid measures of reading comprehension based on the multiple-choice scores and reading comprehension combined with writing based on the short-answer scores. The qualitative data provided specific skills possessed by individual students and if there was need for remediation or the presence of a learning disability. The TBLTB was found to be a reliable and valid assessment instrument in assessing incoming college students' academic literacy skills. It not only assessed reading comprehension skills, but also writing abilities that play an integral part in the academic success of college students.

Although limited research was done on the TBLTB, two important characteristics of the TBLTB made it a highly effective assessment instrument. The reading passage for the TBLTB was taken from a current college psychology textbook. Reading the passage closely represented a true reading assignment that college students might have to do in a freshman college class (Sammons & Davey, 1994). Using an actual college textbook made the task more authentic for students. Thus, the TBLTB made it possible to predict if students would be able to perform those tasks again when reading texts in their college courses.

Secondly, the TBLTB included a writing sample as part of the assessment activity. Writing is a task that college students have to perform every day in college-level courses,

however, very few assessment instruments include a writing component. The writing portion of the test gave added information about students' literacy abilities. First, a writing component can assess more critical thinking skills, such as summarizing and synthesizing information. In multiple-choice items, it is difficult to include questions that go beyond the literal level. In addition, students are required to create an answer. Not only can their thinking processes be evaluated, but also their writing abilities. This added information can be beneficial in the placement of students in college classes and determining if they need an intervention because of underpreparedness or if they are ready for college-level courses.

A Comprehensive Assessment/Placement Plan

There is a diversity of assessment practices employed among community colleges. The instruments that are being used by the community colleges and the ways that the results are interpreted differ greatly from institution to institution. Thus, a composite picture of community college assessment practices was not easy to create. However, it appears that community colleges agree that using only one assessment instrument to assess incoming students is insufficient in providing adequate information about the basic academic literacy abilities and needs of students. In the case of incoming students, an assessment instrument should be able to screen students, diagnose their individual needs, place them in appropriate course work and predict their academic success. In reality no single test is able to do all of those things. Unfortunately, many community colleges use the results of one single literacy test score to guide them in providing services to

incoming students. If the goal of assessing incoming students was designed to assist students in succeeding in college, some changes may need to be made in terms of assessment.

A variety of recommendations have been made to assist community colleges in assessing their incoming students literacy skills more effectively. One popular recommendation is the use of multiple tests (Morante, 1989; Prus & Johnston, 1994; Sawyer, 1990) so a more complete picture of incoming community college students could be developed. Others have recommended a comprehensive assessment plan that includes a variety of factors that vary from plan to plan, but all include components that increase the validity of the assessment process so that more students can be successful in college and ultimately graduate.

Morante (1989) recommended the use of a placement test that measures proficiency rather than aptitude. Aptitude tests can only provide indications of potential problems. They do not provide sufficient information to give an in-depth analysis of students' individual needs. Overall guidelines in choosing an appropriate placement instrument include the tests' content, referencing, discrimination, speededness, reliability, validity and cost, control for guessing and the availability of alternate forms. The reading component of the placement instrument should have the following characteristics: (1) reading passages that are realistic and holistic covering a wide range of subjects, (2) vocabulary assessed in context, (3) criterion-referenced, (4) sufficient discriminatory

power, especially at the lower ends of the proficiency continuum, (5) power rather than timed testing, (6) reliability, and (7) validity.

In addition, Morante suggested that cut ranges should be used in place of cut scores to allow flexibility and latitude in placement decisions. When other factors, such as high school data, age, years away from school, motivation, are correlated with basic academic literacy skills, then appropriate decisions can be made about placement into college-level course work or developmental courses. Developmental courses were included in the process as an essential component for a successful admissions process. They also added mandatory assessment and placement and demonstrated the effect of the program in increasing student success and retention in college.

Sawyer (1990) concurred with Morante that the admissions' process encompasses both the identification of skills and deficiencies of incoming students through the use of a reliable assessment instrument, plus the remediation of students through developmental courses. He emphasized the importance of placing students into courses that matched their instructional levels and how that was even more crucial with underprepared students. Students with adequate skills, background knowledge and cultural knowledge of college are most likely to succeed, but students with deficiencies in basic skills had more disadvantages that might affect their success in college. That is why having a comprehensive assessment and related developmental assistance program are so crucial for underprepared students.

Prus and Johnson (1994) admitted that it was near impossible to find the ideal assessment program, but educators have the responsibility to work towards creating one. According to Prus and Johnson, the ideal assessment program should emphasize validity. Validity is made up of relevance, accuracy and utility. If the assessment methodology does not measure the program and its reliability, or cannot identify the strengths and weaknesses of the program, then it is not effective. It is crucial to use a multimethod and multisource approach to assessment in order to obtain maximum validity.

In conclusion, a review of the literature relevant to the issue of a reliable and valid reading assessment instrument used at a particular community college, led to some conclusions in a number of areas: (1) who is the typical college student of today, (2) the current reading assessment instruments used, and (3) highlights of a better plan.

The typical college student of today is highly diverse academically, culturally, linguistically and socially. The area that higher education can influence more readily is academic diversity, by first assessing incoming students' basic academic skills and then optimizing the fit between students' academic preparation and the appropriate level of instruction.

However, finding a reliable and valid assessment instrument that can identify incoming students' basic academic skills and deficiencies is problematic. Community colleges use a variety of instruments and too many use only one instrument to make placement decisions for their incoming students. Recommendations for the use of a combination of assessments was made, or the institution of a comprehensive assessment

plan that included a variety of measures and services to ensure the success of students, and more importantly, the success of underprepared students.

The particular literacy assessment instruments that were used in this study were investigated for reliability, extent of agreement and concurrent validity. Two of the tests, the NDRT and the DRP, widely used standardized reading instruments were found to be problematic in diagnosing students' individual literacy strengths and weaknesses. A third test, the TBLTB, an informal criterion-referenced test, offered additional information about students, as the reading passage was taken from a college-level textbook and included a writing sample. The suggestion that a variety of assessments be used in assessing incoming students' basic academic literacy skills be used to create a truer picture of students, seems to be warranted.

It was concluded that assessment is a vital component in assisting students to be successful in college. Unfortunately, many colleges are using ineffective instruments and making poor placement decisions about their students' academic preparation for college. More work is needed in this area.

CHAPTER 3

PROCEDURES

This chapter describes and discusses the procedures used to examine three college-level reading assessment instruments used to evaluate academic literacy skills of incoming students in a southwestern community college (SWCC) for the purpose of placing them in developmental reading courses. Sections included in this chapter are: (1) the Southwestern Community College, (2) Community College 1--the Study Site, (3) the Sample, (4) the Developmental Reading Program at SWCC, (5) procedures for gathering data, and (6) the Reading Assessment Instruments.

The Southwestern Community College

SWCC, located in the southwest part of the country, was chosen as the site for this study. SWCC is a two-year institution offering some 2000 courses each semester toward university transfer programs, occupational and developmental education, and general studies courses. Some 53,000 credit and noncredit students take classes on the five separate campuses that are scattered throughout all parts of the city.

Over the last 15 years, the minority student enrollment at SWCC has increased from 24% to 36%, with the white/other group decreasing by some 11%. Native American, African American and Asian American students have remained around the 4% mark, while Hispanic students have nearly doubled in numbers since 1980. Hispanics made up 28% of the student population in Fall 1997.

Women made up the majority of SWCC enrollment comprising 55% of the student population. This percentage has remained quite stable over the last seven years. Between 1985 and 1992, a dramatic shift in gender occurred when the percentage of men declined from 49% to 45%.

Seventy-seven percent of the students were high school graduates, and 27% of those students have graduated from high schools that are within the county where SWCC is located. Of the remaining 16%, 7% of the students have earned a GED and 16% are non-graduates.

The average age of SWCC students during the Fall 1997 semester was 28 years with a median age of 24 years. These statistics have remained consistent since 1993. The proportion of students in their twenties (20-29 years) was 45%, while one-fifth of the enrollment was under 20 years of age. Seventeen percent of the students were 40 years or older.

A little over one-fourth (26%) of its students attended SWCC full-time in the Fall 1997 semester. Full-time enrollment is defined as taking 12 or more semester credit hours. If a student takes courses at two or more campuses which add up to 12 or more units, then that student is considered full-time at the district level, but part-time at each campus.

Most of the SWCC students were state residents (94%) and 87% of those students resided within the county where SWCC is located. These percentages have remained quite stable over the last five years.

In summary, the student enrollment pattern of Fall 1997 was similar to prior years with respect to ethnicity, gender, age, full-time/part-time status and residency. One enrollment statistic that has changed over the last five years is the percentage of new students to higher education. New student status is defined as first time college attendance. New students increased by 16% in just the year preceding 1997. It may be that this increase in new student enrollment is because SWCC is doing a better job of recruiting new students from the local high schools.

A longitudinal study done from 1988 to 1992, conducted by SWCC, supports this supposition as delineated by the following results:

1. The proportion of students composing each minority subgroup equaled or exceeded the corresponding proportion of college age residents in the county. Also, enrollment of the minorities in the 1988 cohort remained relatively stable over time.
2. Both minority and non-minority students were retained at approximately the same retention rates. Retention rates ranged from 50% after the first semester to 14% four years later.
3. Within a five-year period, approximately one-fourth of the students were still enrolled, had graduated or had transferred.
4. Within a five-year period, 6% of full-time students and 3% of all students received a degree or certificate from SWCC. There were few differences between total minority students and non-minority students in terms of graduation rates.

5. The transfer rate to the local university averaged 7% after four years with non-minority students at a somewhat higher rate than minority students.

The Community College 1--The Study Site

Data for this study were collected at only one of the five campuses of SWCC.

This campus (CCI) is located in the far southwest side of the city. CCI offered a wide range of courses serving some 3100 students in credit courses in the Fall 1997 semester, making it the smallest campus in student numbers among the SWCC campuses. The dominant ethnic group attending CCI in 1997 was the Hispanic population which has continued to grow steadily since its doors were opened in 1985. The Hispanic population of CCI has more than doubled since Fall 1991. In the Fall of 1997, 58% of the students were Hispanic which was more than double the Hispanic population for all of SWCC at this time. This was due to the location of CCI and the marketing efforts of that campus to reach out to the surrounding high schools encouraging students to come to college once they graduate. Another outreach strategy that was implemented the year prior to 1997 involved inviting fifth graders of near-by elementary schools to CCI for a special orientation day.

Native Americans are another growing population at CCI. CCI is located on a reservation and great effort has been made to encourage the tribal population to attend. Six percent of CCI's student population was represented by Native Americans during Fall 1997. The African American and Asian American student population at CCI was in line with the rest of SWCC at about 3-4% in 1997.

Sixty-four percent of the students were female on CCI in 1997, which was 9% greater than that for the rest of SWCC. In addition, a larger portion of the students were new students to higher education in 1997. In the Fall of 1997, nearly 10% more (21%) of the students were new enrollees, while 47% were continuing students. The remaining students were transfer students or students returning to SWCC after an absence of one or more semesters. Again, the predominant age range was 20-29 years or 43% of enrollment, with an average age of 27 and a median age of 24. Twenty-four percent of the students were under the age of 20.

The Sample

The subjects included in the study were predominantly freshmen taking classes in 4 of the 15 sections of developmental reading courses offered at CC1 during the 1997 Fall semester. These students were identified as underprepared by the institution, based on their performance on the Nelson-Denny Reading Test (NDRT). The students who participated in this study were all those enrolled in four developmental reading classes taught by the researcher during that semester. Because the students in the sample of this study were all enrolled in intact classes, these classes may not have been equivalent to the other 12 classes at CC1. All of the students in the sample had been advised to take one of several developmental reading classes. While assessment results were primarily employed to determine their developmental class placement, the reasons that they took a particular class may have been due to a number of other factors, including scheduling constraints, transportation issues, or holding a full or part time job. Even though the

students in the sample were not randomly assigned to classes from the larger pool of CCI students, it is probable that they were somewhat representative of the larger group of students taking developmental reading courses during the semester.

The characteristics of the subjects included in the study were different in some ways from the typical college students at SWCC. As was expected, the dominant ethnic group of the sample was Hispanic, comprising 78% of the sample. CCI is located in the southwest part of the city where the majority of Hispanics live. They apparently felt comfortable coming to the local community college in their neighborhood. The Native American portion of the sample was higher also, at 8%, greater than the entire college and CCI. Again, the campus is located on a reservation, and it appears that Native American students felt comfortable attending CCI.

Females made up the larger percentage of the sample numbering 72% in 1997. And surprisingly, 74% of the sample were attending school full time. Seventy-two percent of the sample were first time college students and 90% had graduated from high school. In addition, they were much younger than the typical student at SWCC in 1997, with more than 90% of the students in the sample under the age of 30 having a median age of 21 years of age. Table 3.1 depicts some of the characteristics of the students in the study.

In order to comply with the University of Arizona's Committee to Protect Human Subjects' guidelines, students agreeing to participate in the study signed a Participant Consent Form (See Appendix A). It was made clear to the students that participation was

Table 3.1

Characteristics of the Sample Compared to SWCC and CCI in Fall 1997

CHARACTERISTIC	SWCC	CCI	SAMPLE
Hispanic	28%	61%	78%
Native American	3%	6%	8%
African American	4%	3%	3%
Asian American	4%	2%	3%
White/Other	61%	28%	8%
Male	45%	33%	28%
Female	55%	67%	72%
Full-Time	26%	11%	74%
Part-Time	74%	89%	26%
New Student	21%	22%	72%
H.S. Graduate	75%	77%	90%
Median Age	24	24	21

strictly optional and that their decision would not be in anyway related to course grades.

All students agreed to participate except one student who dropped out of school early in the semester.

The Developmental Reading Curriculum

SWCC reading faculty have recently redesigned its reading curriculum. Wanting “to provide learning opportunities for students who require additional linguistic and/or educational preparation” (the College’s Mission Statement), they have created a comprehensive program not only to assist students in passing reading courses but also to

develop effective academic literacy strategies that would assist them in passing college-level courses and eventually graduate from college.

The curriculum is based on transactional reading theory (Rosenblatt, 1978) that holds that reading primarily involves the construction of meaning and is more than translating the sounds and graphic symbols from the printed page. Reading is taught as a transaction between the reader and the text, while the reader is continually constructing new meaning. The task of the reading classes is to offer a variety of literacy events in which those transactions can take place. After successfully completing the developmental Reading program, students will be able to do the following:

1. Read independently for a variety of purposes.
2. Engage in reading using a critical thinking, problem-solving approach.
3. Value reading as a means of lifelong learning and pleasure.

The curriculum represents a logical progression in the instruction and reading experiences provided for students. There are three reading courses in the sequence leading up to the College Reading Course (Reading 112).

The first course is Reading Fundamentals or Reading 071. This course strives to assist students in developing a sense of themselves as fluent readers and more importantly to help them view reading as a way to derive pleasure from reading. Providing a variety of texts and the opportunities for students to choose reading materials according to their interests, culture and life experiences, it is intended that this will help students to develop a joy for reading, if they do not already feel that way. The course is designed to help

students enhance their background knowledge through reading, become acquainted with basic text features, and develop basic comprehension and vocabulary acquisition strategies. Central to the focus of this course is to provide a community of readers and learners so that students will experience the value of interacting with others as they read and learn.

The next level is Reading Improvement I or Reading 081. This course strives for students to become strategic readers, setting their own purposes for reading, being aware of text structure, and knowing what strategies to use to comprehend new text. Emphasis continues to be placed on literal and text-based comprehension and the acquisition of vocabulary. Basic study strategies and how to use them in content classes are also incorporated in this class. Continuing the use of a community of learners is also included so students will develop good listening, speaking, reading and writing skills in the college classroom as they interact in small cooperative learning groups.

Reading Improvement II or Reading 091 is the final developmental reading class before students take Reading 112 or the college reading course. By this time, students should be reading at or near the twelfth grade level. They are becoming more proficient in interacting with a variety of texts to construct appropriate meanings. A focus on literal comprehension is continued but more emphasis is placed on developing critical comprehension strategies. These comprehension strategies are used on a variety of texts. Ultimately, students will use them independently as they read and interact with other college level texts.

Not all students take all three of the developmental reading courses. Depending on how they score on the Nelson-Denny Reading Test (NDRT) they may be advised to go immediately to the second or third level classes. Also, an instructor can recommend that a student skip a course along the way through the three-course sequence.

Procedures for Gathering Data

All students interested in taking courses at the SWCC are first required to take placement tests in reading, writing and math. In the Fall semester 1997, the Nelson-Denny Reading Test (NDRT) Form E (Brown, Bennet & Hanna, 1981) was used as the reading assessment instrument to evaluate students' academic reading skills and place them in an appropriate developmental reading class.

The students went to the testing center of CC1, which provided a quiet area for testing. The students checked in with the coordinator who provided them with a tape recorder and the NDRT test packet. The tape recorder contained a recording that gave each student a set of oral instructions on how to complete the test and enabled them to time themselves. This was to assure that all students were treated in the same manner and the guidelines for administering the NDRT were followed in a manner consistent with the NDRT administrator's manual.

The first part of the NDRT determines the student's reading rate by asking students to count the number of words that they are able to read during the first five minutes of the test. Then the NDRT assesses students' vocabulary and comprehension skills. The first part, the vocabulary subtest, consists of 100 multiple-choice items which

require students to choose the best definition of a particular word from five options. Fifteen minutes are given to complete the vocabulary subtest. Once the time is up, students are notified to stop working on that part of the test and given instructions for the next subtest, comprehension.

The comprehension subtest contains eight short passages drawn from the humanities, the natural and biological sciences, and the social sciences followed by 36 multiple-choice comprehension questions. Total score possibilities on the NDRT test are 172 points, with one point for each correctly answered vocabulary question and two points for each correctly answered comprehension question. Table 3.2 shows the NDRT cut off scores for placement into the developmental reading courses and the equivalent grade level.

Table 3.2

Criterion NDRT Raw Scores for Placement into Four SWCC Reading Courses

TOTAL RAW SCORE ON NDRT	GRADE LEVEL EQUIVALENCY	RECOMMENDED READING COURSE
Below 50	5.0 - 7.5	READING 071
50-64	7.8 - 9.9	READING 081
65-76	10.1 - 11.5	READING 091
77-85	11.6 - 12.6	READING 112

The testing center entered the score of the NDRT into the computer system which contained all vital information about each student. Students were then directed to see an

advisor who helped them determine their course of study. The advisor could only recommend that students take the course or courses that the assessment instrument indicated. Placement was not mandatory. If a student did not want to take the course that the NDRT results recommended, that student was not required to do so. Many students also had time constraints because of work or family, so they often enrolled in the courses that best fitted their schedules. Also, not all students sat down with an advisor to work out their course schedules. Many students slipped through the cracks and never took the assessments or met with an advisor.

On the first day of classes, students electing to take a developmental reading course were again evaluated on their reading abilities with the Degrees of Reading Power (DRP) another standardized reading test that is widely used in some colleges to determine students' reading skills (The College Board, 1980). This test was administered in the classroom by the classroom teacher. Faculty were instructed to give the same directions to each class, but whether or not that was done properly is unknown.

Table 3.3 shows the cut off scores for placement into the reading courses.

If the total raw score on the DRP showed that a student should be placed in a course that was two levels above or below the placement suggestion from the NDRT, then the student was advised to change to the class that the DRP recommended. However, it was found that many students were unable to change classes at that time in the semester due to work and class schedules. So, they either stayed in the class

Table 3.3**Criterion DRP Raw Scores for Placement into Four SWCC Reading Courses**

TOTAL RAW SCORE ON DRP	RECOMMENDED READING COURSE
50 and Below	READING 071
51-57	READING 081
58-61	READING 091
62 and Above	READING 112

recommended by the NDRT, changed to the class recommended by the DRP or dropped the reading class to take it at another time.

During the second week of classes, the Turrentine Bradley Literacy Testing Battery, Form A (TBLTB) (Turrentine & Bradley, 1997) was administered to the students in the four developmental reading classes that were taught by the researcher. The TBLTB was created by P. Turrentine and J. Bradley at the University of Arizona (see Dreams Reading Passage and test questions in Appendix B). The test required students to read a six-page passage from a freshmen psychology textbook (see permission letter in Appendix C), which represented the level of reading that is required by most college courses. After students silently read the passage, they were asked to complete 10 multiple-choice and 11 short-answer questions that were tied to the passage. The TBLTB evaluated how students extracted concepts from a lengthy text and then how they expressed those concepts in writing.

The authors of the TBLTB did not establish cut scores other than determining that if students answered seven of the multiple-choice questions correctly and answered seven of the short-answer questions correctly, they passed the tests. This researcher followed that criteria and added the breakdown of scores in Table 3.4 to show how students might be placed into the developmental reading courses depending on the raw scores they earned on both the short-answer and multiple-choice sections of the TBLTB.

Table 3.4

Criterion TBLTB Raw Scores for Placement into Four SWCC Reading Courses

RAW SCORE ON MC	RAW SCORE ON SA	RECOMMENDED READING COURSE
0-2	0-2	READING 071
3-4	3-4	READING 081
5-6	5-6	READING 091
7 and Above	7 and Above	READING 112

In addition to assessing students' literacy skills by three different assessment instruments, a short survey inquiring about age, sex, ethnicity, first language and place of birth was completed by the students in the sample. Educational background data were also gathered, such as high school attended, graduation status and placement in high school bilingual or ESL classes. Each subject's enrollment status was gathered, such as, if the subject was a first time student, a continuing student, or a student returning to college after being out of school for more than a semester. Finally, subjects were asked if

they were working while going to school and if working, to indicate how many hours per week that they did work. This information was gathered so that a better profile about the sample population could be developed. Research has found that many non-academic factors come into play when determining the success of students in particular classes, programs, or college in general (Astin, 1991; Tinto, 1987, 1988). Looking at just scores on the three assessment instruments did not seem sufficient to determine if students had been assessed and placed appropriately in college courses.

The NDRT was scored by the testing center, and the researcher was given a computer printout that listed each student and the scores that the students attained in terms of both pre- and post-tests scores. The DRP was scored by the college's reading lab coordinator. It is assumed that she graded all DRP tests correctly.

The TBLTB was graded by the researcher. The multiple choice section had set answers so it is assumed that the researcher graded all tests on this section correctly. To determine the extent of inter-rater reliability on the short answer section, this researcher and the developer of the TBLTB randomly selected the tests of 10 subjects from the sample as a subsample and independently scored the answers of those tests according to the answer key provided with the TBLTB. The extent of agreement between the two raters was 100%. The results of their work are shown in Table 3.5.

The Reading Assessment Instruments

The Nelson-Denny Reading Test (NDRT) is one of the most widely used standardized reading tests in colleges throughout the nation (Wood, 1989; Maxwell,

Table 3.5**Inter-Rater Reliability on the TBLTB-SA Format for 10 Subjects**

Cases	Rater 1	Rater 2	Differences	Pass/Fail*
1	7.5	8.0	-.5	Pass/Pass
2	7.5	7.0	.5	Pass/Pass
3	6.5	5.0	1.5	Fail/Fail
4	6.0	6.5	-.5	Fail/Fail
5	6.0	6.5	-.5	Fail/Fail
6	5.5	5.0	.5	Fail/Fail
7	5.0	4.0	1.0	Fail/Fail
8	3.5	2.5	1.0	Fail/Fail
9	2.5	3.0	-.5	Fail/Fail
10	2.0	1.5	.5	Fail/Fail
Mean	5.2	4.90	0.70	
SD	1.95	2.13	0.35	

* Passing was defined as obtaining a raw score of 7.0 or better

1997). Since its inception in 1929, the NDRT has been revised a number of times. There are eight versions of the test (A & B, C & D, E & F, and the newest versions G & H).

Tests were created in pairs to produce a comparable pre- and post-test system, although all forms of the test follow the same format and supposedly can be used interchangeably.

The NDRT consists of two subtests, vocabulary and comprehension, both of which are timed and written in multiple-choice format. The vocabulary subtest consists of 100 items. Students have 15 minutes to choose the best definition for each word. The

comprehension subtest consists of eight short passages drawn from a variety of subjects, such as the humanities, the natural and biological sciences and the social sciences, that are to be read then respond to 36 multiple-choice questions. Twenty minutes are provided to complete this section. Reading rate is measured by counting the number of words read in one minute and extrapolated.

The total score on the NDRT is obtained by assigning one point for each vocabulary question answered correctly and two points for each comprehension question answered correctly. The maximum possible total score is 172 points. Using total raw scores on the NDRT and the minimum raw scores on the comprehension subtest the following grade equivalency chart was determined (see Table 3.6).

Table 3.6

Grade Equivalency for NDRT-Form E

RANGE OF TOTAL RAW SCORES	GRADE LEVEL EQUIVALENCY	MINIMUM COMPREHENSION RAW SCORE	GRADE LEVEL EQUIVALENCY
86 and above	12.7+	42	12.4
77 - 85	11.6 - 12.6	36	10.7
65 - 76	10.1 - 11.5	32	9.4
50 - 64	7.8 - 9.9	24	6.7
Below 50	5.0 - 7.5		

Authors of the test purport that the test provides, “a trustworthy ranking of student ability in three areas of academic achievement: reading comprehension, vocabulary

development and reading rate” (Brown, Bennet, & Hanna, 1981). They recommend that the test be used to screen and place both superior and remedial or developmental students to predict their success in college classes and to diagnose reading weaknesses (pp. 6-10).

SWCC required all incoming students in 1997 to take the NDRT before registering for courses. The raw score earned on the NDRT determined the courses that students could take, whether they be developmental reading courses or regular college-level courses. For instance, if students scored 86 or above on the NDRT, they were advised that they can take college-level courses since that score indicates that they are reading at a twelfth grade level or higher. If they scored lower than a score of 86, they were advised to take one or more of the developmental reading classes, so they could improve their literacy skills. SWCC only advised students to take these courses in the Fall 1997. There was no mandatory placement. The Reading Department faculty at SWCC is aware of the research that notes students, who come to college underprepared for college-level text, are at a greater risk of dropping out or failing in college than students who are reading at a college level. That is why assessing incoming students’ academic skills is important plus providing adequate advising, so students become aware of both their own deficiencies and the programs that are available to them.

Several years ago, the SWCC reading faculty, not content with the NDRT, decided to pilot another reading test, to be used in conjunction with the NDRT. One major reason they questioned the effectiveness of the NDRT was because SWCC has a very diverse student population. Since the NDRT had been normed on a very different

student population from that of SWCC, they questioned if the NDRT was capable of evaluating the reading skills of SWCC's student population. Because of these concerns, they chose to augment the NDRT with a second test, the Degrees of Reading Power (DRP), another college-level reading instrument widely used by community colleges.

The DRP is a modified-cloze test composed of eight prose passages arranged in an order of increasing difficulty. The passages, taken from the Encyclopedia Britannica, are approximately 350 words in length each and contain modified cloze items which replace words from the passage. The reader is provided with five single word options, from which to choose the best word to select for the cloze items. All of the options are designed to complete the sentence, but only one is correct.

A major difference between the DRP and the NDRT is that the DRP is not timed. Students are allowed to spend as much time as they want completing the test. In addition, three separate levels of comprehension are reported. An independent comprehension level, which indicates the kinds of texts a student can read and comprehend independently. The student uses appropriate strategies on his own when comprehension begins to fade, using both the text and background knowledge, as resources for understanding. The instructional level indicates when a student will experience a need for instructional support to understand the text. And the frustration level indicates the text with which the student will have difficulty reading and comprehending even with instructional support.

The third assessment instrument used in this study was the Turrentine Bradley Literacy Testing Battery (TBLTB), a criterion-referenced, informal reading inventory. The test consists of a six-page reading passage from a freshman psychology text (Gleitman, 1991) about Dreams. Since a large majority of college students are required to take a psychology course, the choice of this subject seemed to be logical. In addition, the TBLTB Dreams passage should be representative of the reading level of most first-year college texts.

Subjects were told to read the passage silently. Once they finished reading, they were directed to raise their hands and the researcher would bring them a set of 10 multiple-choice questions that pertained to the passage for them to answer. The questions were taken from the study guide (Jonides & Rozin, 1991) which accompanied the psychology textbook with the addition of several questions submitted by other doctoral students. The subjects were not timed and were allowed to reread the passage while answering the questions. If a subject was able to answer 70% of the questions correctly, it was assumed that the student was able to comprehend college-level reading material. In essence, it was assumed that they were able to decode, comprehend and identify key concepts in the passage.

The subjects were told to raise their hands when they had completed the multiple-choice questions and that part of the test would be picked up and they would be given a set of 11 short-answer questions. These questions not only addressed the literal level of comprehension but also inferential comprehension, plus writing competency. Again, the

subjects were told there was no time limit, and they were able to refer back to the passage to answer the questions. It was assumed that if subjects passed this portion of the test with at least 70% of the questions answered correctly, they were reading and writing at the college level. If they failed to pass this portion of the test, but passed the multiple-choice section, then it may be assumed that the subjects needed assistance in the area of writing and expressing their ideas. Those who failed both sections of the test probably lacked the necessary literacy skills to perform college-level work.

All three of the above tests were administered to 88 CC1 students in the sample during the Fall semester of 1997.

Analysis of the Data

Correlations between the following tests were determined using Pearson Product-Moment Coefficients:

1. The raw scores on the Nelson-Denny Reading Test Vocabulary Subtest (NDRT-V) with the raw scores on the DRP.
 2. The raw scores on the Nelson-Denny Reading Test Comprehension Subtest (NDRT-C) with the raw scores on the DRP.
 3. The raw scores on the Nelson-Denny Reading Test Total (NDRT-T) with the raw scores on the DRP.
 4. The raw scores on the NDRT-V with the raw scores on the Turrentine/Bradley Literacy Testing Battery Multiple-Choice (TBLTB-MC) Format.
 5. The raw scores on the NDRT-C with the raw scores of the TBLTB-MC.
-

6. The raw scores on the NDRT-T with the raw scores on the TBLTB-MC.
7. The raw scores on the NDRT-V with the raw scores on the Turrentine/Bradley Literacy Testing Battery Short-Answer (TBLTB-SA) Format.
8. The raw scores on the NDRT-C with the raw scores of the TBLTB-SA.
9. The raw scores on the NDRT-T with the raw scores of the TBLTB-SA.
10. The raw scores on the DRP with the raw scores of the TBLTB-MC.
11. The raw scores on the DRP with the raw scores of the TBLTB-SA.
12. The raw scores on the TBLTB-SA with the raw scores on the TBLTB-MC.

Kappas and one-tailed ANOVAs were computed among the scores of the following reading tests to determine concurrent validity, reliability and extent of agreement in placement into developmental reading courses:

1. The raw scores of the NDRT-T with the raw scores of the DRP.
2. The raw scores of the NDRT-T with both the scores of the TBLTB-MC and the TBLTB-SA.
3. The raw scores of the DRP with both the scores of the TBLTB-MC and the TBLTB-SA.
4. The scores of the TBLTB-MC with the scores of the TBLTB-SA.

In addition to the quantitative research done in this study, three students were interviewed to add depth to the study. The interviews took place during the Spring 1998 semester, one semester after the subjects had participated in the study during the Fall, 1997. The three students were individually interviewed and asked to tell about their

perceptions of themselves as students and how they perceived their own literacy skills. They were asked to share their feelings about the tests administered and if they felt they had been placed in the appropriate developmental reading class because of the results of those tests. The subjects told the researcher about the other courses they had taken during the Fall 1997 and the Spring 1998 semesters. They also described how they thought they had performed in those courses.

The qualitative data gathered in those interviews were used in combination with the quantitative results to come to some conclusions about actual students who participated in this study. By combining both quantitative and qualitative information, a better picture might be created of a student going through the assessment and placement process. As educators and administrators know, students are more than a compilation of test scores. They are living, breathing human beings with needs and desires that affect their performance in college. Adding the cultural and personal factors of some students might authenticate the scores earned by the students.

CHAPTER 4

FINDINGS

The purpose of this chapter is to present the findings of the study. The findings and the discussion are presented under the following headings: (1) introduction, (2) description of the analysis procedures, (3) correlational findings, (4) extent of agreement findings, (5) ANOVA findings, and (6) qualitative data.

Introduction

Three college reading assessment instruments, the Nelson-Denny Reading Test (NDRT), the Degrees of Reading Power (DRP) and the Turrentine/Bradley Literacy Testing Battery (TBLTB) were administered to a sample of 88 students at a southwestern community college (SWCC) in the Fall semester of 1997. The purpose of the study was to investigate the interrelationships of these three tests for reliability, concurrent validity and extent of agreement for placement into developmental community college reading courses.

Description of the Analysis Procedure

Correlations were computed between the raw scores of the following tests using the Pearson Product-Moment Correlation Coefficients:

1. The raw scores on the Nelson-Denny Reading Test-Vocabulary Subtest (NDRT-V) with the raw scores on the DRP.
-

2. The raw scores on the Nelson-Denny Reading Test-Comprehension Subtest (NDRT-C) with the raw scores on the DRP.
3. The raw scores on the Nelson-Denny Reading Test-Total (NDRT-T) with the raw scores on the DRP.
4. The raw scores on the NDRT-V with the scores on the Turrentine/Bradley Literacy Testing Battery Multiple-Choice (TBLTB-MC) Format.
5. The raw scores on the NDRT-C with the scores on the TBLTB-MC.
6. The raw scores on the NDRT-T with the scores on the TBLTB-MC.
7. The raw scores on the NDRT-V with the scores of the Turrentine/Bradley Literacy Testing Battery-Short-Answer Format (TBLTB-SA).
8. The raw scores on the NDRT-C with the scores on the TBLTB-SA.
9. The raw scores on the NDRT-T with the scores on the TBLTB-SA.
10. The raw scores on the DRP with the scores on the TBLTB-MC.
11. The raw scores on the DRP with the scores on the TBLTB-SA.
12. The raw scores of the TBLTB-SA with the raw scores of the TBLTB-MC.

Kappas and one-tailed ANOVAs were computed among the scores of the following reading tests to determine concurrent validity, reliability and extent of agreement in placement into developmental reading courses:

1. The raw scores of the NDRT-T with the raw scores of the DRP.
 2. The raw scores of the NDRT-T with both the scores of the TBLTB-MC and the TBLTB-SA.
-

3. The raw scores of the DRP with both the scores of the TBLTB-MC and the TBLTB-SA.
4. The scores of the TBLTB-MC with the scores of the TBLTB-SA.

Explanation of the Kappa

Kappa is a coefficient appropriate for the measurement of extent of agreement between independent predictors for data expressed in nominal data scales (Cohen, 1960).

In the measurement of extent of agreement, there are two relevant quantities:

P_o = proportion of obtained agreement

P_c = proportion of expected chance agreement

The coefficient of Kappa is the proportion of agreement after chance agreement has been removed. The mathematical equation for determining Kappa is:

$$K = \frac{P_o - P_c}{N - P_c}$$

P_o	=	agreement observed
P_c	=	chance agreement
N	=	number of cases

Z value at .001 significance = 2.326 for a one-tailed test.

Correlational Findings

Correlations were calculated for pairings of the three tests and their subtests: the NDRT-V, NDRT-C and the NDRT-T; the DRP and the TBLTB, both MC and SA format. Table 4.1 shows the correlation coefficients among the three tests.

All of the obtained correlations were significant at the .01 level, ranging from a low of .3435 to a high of .8461. The correlation coefficients were significantly different from 0, which means that there were positive linear correlations among all of the tests,

Table 4.1

Pearson Product-Moment Correlation Coefficients Between the Scores of the Tests

TESTS	NDRT-V	NDRT-C	NDRT-T	DRP	TBLTB-MC	TBLTB-SA
NDRT-V	---	.3602	.7961	.3746	.3435	.3800
NDRT-C		---	.8461	.4980	.3860	.3720
NDRT-T			---	.5332	.4453	.4732
DRP				---	.3506	.4681
TBLTB-MC					---	.3703

but many of them were moderately low. We cannot predict with accuracy a student's score on one test, if we know his/her score on another test. However, since there were positive linear correlations among all the tests, we can predict that scores on one test will increase as the scores on the other tests increased.

The higher obtained correlations were between the NDRT-V and the NDRT-T ($r = .7961$) and the NDRT-C and the NDRT-T ($r = .8461$). One would expect a higher correlation in the case of the NDRT-T and its two subtests, NDRT-V and NDRT-C because of the whole-part phenomenon. The obtained correlations show that the two NDRT subtests (V and C) are contributing uniquely to the NDRT-T, making the test reliable and valid, the way a test should be constructed.

The correlation between the NDRT-T and the DRP was low ($r = .5332$). This was surprising given the fact that both of the tests were designed to assess literacy competencies. While the two tests are positively related, they appear to be somewhat

different in what they measure. The moderately low correlation may have been attributed to the following differences in the tests: (1) format, (2) timed, or (3) location of the test-taking activity. The DRP uses a modified cloze format embedded in short reading passages. Also, the passages are arranged in an order of increasing difficulty. The NDRT uses multiple-choice items for assessing isolated vocabulary words and short reading passages in no particular order of difficulty. The format of the tests may have contributed to the differences. In addition, the NDRT has strict time limits on both subtests, while the DRP allows students to take as much time as they wish to complete the test.

Another possible factor contributing to the low correlation between the two tests was the fact that both tests were administered in very different environments. Incoming students were directed to the testing center to take three placement tests before they were able to register for classes, and the NDRT was the literacy assessment instrument. Students were new to the campus, alone and probably apprehensive about the notion of taking tests if they didn't know in advance that they had to take placement exams. The stress factor in this situation might have contributed to the scores earned on the NDRT. In contrast, the DRP was administered in the classroom by the instructor. Granted it was the first day of class and students may have been apprehensive also, but the teacher was there to give personal instructions in taking the test, rather than a tape recorder that was used for the NDRT. Students may have felt more comfortable in this environment.

The correlation between the DRP and the NDRT-C ($r = .4980$) was higher than the correlation between the DRP and the NDRT-V ($r = .3746$). This may be due to the

fact that the DRP is more like the format of the NDRT-C than the NDRT-V. Both the DRP and the NDRT-C have short reading passages with multiple-choice items to complete. The NDRT-V assesses vocabulary in isolation in the multiple-choice format, while the DRP does not directly measure vocabulary, which might explain the differences.

The moderately low correlation between the TBLTB-MC and the TBLTB-SA ($r = .3703$) was surprising because both of the tests had parallel questions that were based on the same reading passage about Dreams. There are two possible reasons for the relatively low correlation : (1) the relatively few test items on the TBLTB and (2) a floor effect. In the case of the TBLTB, it was apparent that the test was too hard for most of the students because virtually all of them failed the test. Most of the scores were at the lower end of the continuum, demonstrating little variation. When variability is low, relatively low correlations usually are the result.

In addition, it was surprising that the TBLTB-SA correlated higher with the DRP ($r = .4681$) and the NDRT-T ($r = .4732$) than the TBLTB-MC, the DRP ($r = .3506$) and the NDRT-T ($r = .4453$). The TBLTB-SA had a very different format, writing short-answer responses to questions. One reason that might explain these differences was the fact that the students completed the TBLTB-MC section first. The students had already read the passage one time and had skimmed through the passage several times looking for answers to the multiple-choice questions. This may have helped some of them to get higher scores on the TBLTB-SA which in turn increased the extent of score variation on

the TBLTB-SA. Possibly, the second look at the reading passage, or maybe a more focused reading of the passage, improved their scores to the point that the test variance was increased, thereby slightly increasing the correlations.

Extent of Agreement Findings

All incoming students interested in taking courses at a southwestern community college (SWCC) were required to take placement tests in reading, writing and mathematics before attending classes. In the Fall semester of 1997, the Nelson-Denny Reading Test (NDRT), Form E, was used as the reading assessment instrument to evaluate students' academic literacy skills. The NDRT is a timed, standardized, norm-referenced test that evaluates vocabulary and comprehension in a multiple-choice format. Depending on the scores earned, students were placed in the following three developmental reading classes: Reading 071, Reading 081 or Reading 091, as well as Reading 112, a college-level course. Reading 071 was the lowest developmental reading class, and students with NDRT raw scores below 50 were initially placed into that course. Placement into Reading 071 meant that a student needed basic literacy skills instruction. Students earning NDRT raw scores between 50 and 64 were placed into Reading 081. Those students placed into Reading 081 were thought to be somewhat better prepared than the students placed in Reading 071. Students in Reading 081 were given instruction in developing comprehension reading strategies. Students were placed into Reading 091 when they obtained NDRT raw scores between 65 and 76. Instruction in Reading 091 involved developing critical comprehension strategies. Those students who scored above

77 on the NDRT were placed into Reading 112, a college-level reading course. Reading 112 was not considered to be a developmental course. When analyzing the placement of students into the three developmental reading classes, the NDRT was used as the comparison measure.

On the first day of classes, students enrolled in the above three developmental reading classes were again evaluated with the Degrees of Reading Power (DRP), an untimed, standardized, modified cloze test. Table 4.2 depicts the number of times the NDRT and the DRP placed students in the same developmental reading classes, and the number of times they differed in their placement decisions.

Table 4.2

The Number of Students Correctly Placed, Underplaced and Overplaced by the DRP When Compared with the NDRT

		NDRT		
Courses		071	081	091
DRP	071	22	16	5
	081	13	6	6
	091	2	7	6
	112		2	3

Of the 88 students in the sample, who were evaluated by both the NDRT and the DRP, there was agreement in the placement of 34 of the students, or 39% of the sample. The two tests agreed the greatest in the placement of 22 students into Reading 071, the lowest level reading class. The tests agreed on placing six students into Reading 081 and six students into Reading 091 for a total of 34 students for whom agreement was obtained between the NDRT and the DRP. The 54 remaining students, or 61% of the sample were placed in different classes. The DRP underplaced 27 students into lower classes than the NDRT, placing 22 students one class level lower and five students two class levels lower. Even though the NDRT has a track record for being a hard test, for these students it was actually easier, and they were placed into a higher level developmental reading class than the level the DRP recommended. The DRP also overplaced 27 students when compared to the NDRT, placing 23 students one class higher and four students two classes higher. In conclusion, the NDRT and the DRP demonstrated low to moderate agreement in placing incoming students into the three different developmental reading courses.

During the second week of classes the Turrentine/Bradley Literacy Testing Battery (TBLTB), Form A was administered to the students enrolled in the three developmental reading classes. The TBLTB is an informal criterion-referenced test based on a six-page reading passage taken from a university psychology text. Students were required to read the passage silently then answer both short-answer and multiple-choice questions related to the passage. Each part of the test, both the multiple-choice and the short-answer

section, was considered a separate test when comparisons were made among tests on placement decisions.

Table 4.3 depicts the number of times the NDRT and the TBLTB-MC placed students in the same developmental reading classes and the number of times they differed in their placement decisions.

A total of 85 students were evaluated with the TBLTB-MC and the NDRT. There was agreement in the placement of 33 students, or 39% of the sample. The two tests agreed the greatest in the placement of 15 students into Reading 081, the second level developmental reading class. The tests agreed on placing 13 students into Reading 071

Table 4.3

The Number of Students Correctly Placed, Underplaced and Overplaced by the NDRT When Compared with the TBLTB-MC

		NDRT		
Courses		071	081	091
TBLTB-MC	071	13	4	2
	081	6	15	4
	091	6	8	5
	112	4	10	8

and five students into Reading 091. The remaining 52 students or 61% of the sample were placed into different classes by the two tests. The TBLTB-MC underplaced ten students into lower classes than the NDRT, placing eight students one class level lower and two students two class levels lower. The TBLTB-MC overplaced 42 students when compared with the NDRT: 22 students one class level higher, 16 students two class levels higher and four students three class levels higher. The TBLTB-MC results indicated that 22 students did not need a developmental reading course. In conclusion, the NDRT and the TBLTB-MC demonstrated low to moderate agreement in placing incoming students into the three developmental reading courses and the college-level reading course.

A total of 84 students were evaluated by the NDRT and the TBLTB-SA (see Table 4.4). There was agreement in placement of 24 students, or 28% of the sample. The

Table 4.4

The Number of Students Correctly Placed, Underplaced and Overplaced by the NDRT When Compared with the TBLTB-SA

		NDRT		
Courses		071	081	091
TBLTB-SA	071	9	4	0
	081	9	8	8
	091	10	18	7
	112	1	6	4

two tests agreed that nine, eight, and seven students should be placed into Reading 071, Reading 081 and Reading 091 respectively. The remaining 60 students or 72% of the sample were placed into different classes by the two tests. The majority of the students (40) were overplaced by the TBLTB-SA. When compared to the NDRT, 31 students were placed one class level higher by the TBLTB-SA, 16 students two class levels higher and one student three class levels higher. The TBLTB-SA indicated that 11 students demonstrated college-level literacy skills and did not need a developmental reading course. In conclusion, the NDRT and the TBLTB-SA demonstrated low agreement in placing incoming students into the three developmental reading courses and the college-level reading course.

A total of 85 students were evaluated by the DRP and the TBLTB-MC see Table 4.5). There was agreement in placement of 22 students, or 25% of the sample. The two tests agreed the greatest in the placement of 12 students into Reading 071, the lowest level developmental reading class. The tests agreed on the placement of three students into Reading 081, five students into Reading 091 and two students into Reading 112. The remaining 63 students or 75% of the sample were placed differently by the two tests. The TBLTB-MC underplaced 12 students into lower classes than the DRP. The TBLTB-MC overplaced 52 students, placing 30 students one class level higher than the DRP, 16 students two class levels higher than the DRP and six students three class levels higher.

Table 4.5

**The Number of Students Correctly Placed, Underplaced and Overplaced by the
DRP When Compared with the TBLTB-MC**

		DRP			
Courses		071	081	091	112
TBLTB-MC	071	12	4	1	2
	081	19	3	4	
	091	8	5	5	1
	112	6	8	6	2

Both the DRP and the TBLTB-MC tests agreed that two of the students did not need a developmental reading course. In conclusion, the DRP and the TBLTB-MC demonstrated low agreement in placing incoming students into the three different developmental reading courses and the college-level reading course.

A total of 84 students were evaluated by the DRP and the TBLTB-SA (see Table 4.6). There was agreement in the placement of 18 students or 21% of the sample. The two tests agreed the greatest in the placement of nine students into Reading 071, the lowest developmental reading class. The tests agreed on placement of two students into Reading 081 and seven students into Reading 091. The remaining 66 students or 79% of the sample were placed into different classes by the two tests. The TBLTB-SA

Table 4.6

**The Number of Students Correctly Placed, Underplaced and Overplaced by the
DRP When Compared with the TBLTB-SA**

		DRP			
Courses		071	081	091	112
TBLTB-SA	071	9	4		
	081	18	2	2	3
	091	16	11	7	1
	112	2	3	6	

underplaced 10 students into lower classes than the DRP, seven students into one class level lower and three students two class levels lower. The TBLTB-SA overplaced 56 students, placing 35 students one class level higher than the DRP, 19 students two class levels higher than the DRP and two students three class levels higher than the DRP. In conclusion, the DRP and the TBLTB-SA demonstrated low agreement in the placement of incoming students into the three different developmental reading courses and the college-level reading course.

Of the 84 students evaluated by both the TBLTB-MC and the TBLTB-SA, there was agreement in the placement of 29 students, or 35% of the sample (see Table 4.7).

Table 4.7

**The Number of Students Correctly Placed, Underplaced and Overplaced by the
TBLTB-MC When Compared with the TBLTB-SA**

		TBLTB-MC			
	Courses	071	081	091	112
RBLTB-SA	071	6	3	1	3
	081	7	11	4	3
	091	5	10	8	12
	112	1	6	4	

The two tests agreed the greatest in the placement of 11 students into Reading 081. The tests agreed on the placement of six students into Reading 071, eight students into Reading 091 and four students into Reading 112. The remaining 55 students or 65% of the sample were placed in different classes by the two tests. The TBLTB-SA underplaced 26 students into lower classes than the TBLTB-MC. The TBLTB-SA overplaced 33 students into higher classes than the TBLTB-MC. Although 22 students passed the TBLTB-MC with 70% accuracy or better, and 11 students passed the TBLTB-SA with 70% accuracy or better, only four students passed both sections of the test demonstrating they did not need a developmental reading course. In conclusion, the TBLTB-MC and the TBLTB-SA demonstrated low to moderate agreement in the placement of incoming

students into the three developmental reading courses and the college-level reading course.

A statistic that may be used to determine the significance of the extent of agreement between two tests is Kappa (K). Kappas were computed between pairings of all the tests and the results are found in Table 4.8.

Table 4.8

Kappa Coefficients Between the Tests.

TESTS	Kappa	Std Error	Z-Value	Sign @ .01
NDRT:DRP	.07549	.07587	0.99	No
NDRT:TBLTB-MC	.21994	.06991	3.14	Yes
NDRT:TBLTB-SA	.02694	.06613	0.41	No
DRP:TBLTB-MC	.08791	.07017	1.25	No
DRP:TBLTB-SA	.00996	.06353	0.16	No
TBLTBMC:SA	.18670	.06593	2.83	Yes

There was no significant agreement in the placement of students between the NDRT and the DRP ($K = 0.99$).

There was a significant agreement in the placement of students between the NDRT and the TBLTB-MC ($K = 3.14$). Those results were significant at the .01 level.

There was no significant agreement in the placement of students between the NDRT and the TBLTB-SA ($K = 0.41$).

There was no significant agreement in the placement of students between the DRP and the TBLTB-MC ($K = 1.25$).

There was no significant agreement in the placement of students between the DRP and the TBLTB-SA ($K = 0.16$).

There was significant agreement in the placement of students between the multiple-choice format and the short-answer format of the TBLTB ($K = 2.83$). Those results were significant at the .01 level of confidence.

In summary, the Kappa analysis revealed that there were two test combinations that showed agreement regarding student placement into the three developmental reading classes: (1) the NDRT and the TBLTB-MC and (2) the TBLTB-MC and the TBLTB-SA. The extent of agreement between these test pairings was great enough so that it might have occurred by chance only one time out of 100. This means that 99 times out of 100, agreement of this magnitude was due to similarities between the tests and the results that they produced.

It was possible for these tests to have statistical significance in terms of extent of agreement in placement, but not practical significance. In reality, only 39% of the students in the sample were placed in the same developmental reading class by both the NDRT and the TBLTB-MC and only 35% of the students were placed in the same developmental reading class by the TBLTB-MC and the TBLTB-SA. Over 60 percent of the students in each case were placed differently by the tests. Even though the placement decisions by the tests agreed statistically, for practical purposes, it does not appear they

agreed enough when the consequences of either overplacing or underplacing students in developmental reading courses is considered.

ANOVA Findings

Students were tested on the NDRT before they could register for classes during the Fall 1997 semester. Depending on the scores earned on the NDRT, the students were placed into the following developmental reading classes: Reading 071, Reading 081 or Reading 091, as well as reading 112, a college-level reading class. Students were evaluated by two more literacy assessment instruments: The DRP and the TBLTB during the first two weeks of classes. Once students were identified in terms of how they were placed in the three developmental reading courses, the mean scores and standard deviations were computed on how the students scored on the TBLTB. After that, students were identified in terms of how they were placed into the three developmental reading classes by the DRP, then the mean scores and standard deviations were computed on how the students scored on the TBLTB. ANOVAs were computed to determine if given the fact that students were placed in a given class by either the NDRT or the DRP, whether or not the TBLTB mean scores would increase as the difficulty of the course increased.

Tables 4.9, 4.10, and 4.11 identify the number of students who were placed into each class, Reading 071, Reading 081 or Reading 091, the mean scores and standard deviations based on the placement recommendations of the (1) NDRT and (2) the DRP.

Table 4.9

Mean Scores and Standard Deviations Computed According to Placement by the NDRT When Computed with the DRP

	READING 071			READING 081			READING 091		
	n	M	SD	n	M	SD	n	M	SD
NDRT-V	19	21.4	5.72	43	25.3	6.47	26	30.8	7.02
NDRT-C	19	18.7	7.56	43	28.9	6.15	26	33.5	6.63
NDRT-T	19	40.2	9.07	43	54.0	9.24	26	64.3	10.94
DRP	9	41.9	11.97	43	48.8	8.86	26	49.8	10.49

Table 4.10

Mean Scores and Standard Deviations of the TBLTB-MC and TBLTB-SA, According to Placement by the NDRT

	READING 071			READING 081			READING 091		
	n	M	SD	n	M	SD	n	M	SD
TBLTBMC	17	3.58	2.06	41	4.80	2.30	26	5.46	2.35
TBLTB SA	17	3.50	2.74	41	4.96	1.35	26	5.38	2.05

Table 4.11

**Mean Scores and Standard Deviations of the TBLTB-MC and TBLTB-SA,
According to Placement by the DRP**

	READING 071			READING 081			READING 091		
	n	M	SD	n	M	SD	n	M	SD
TBLTBMC	45	4.07	2.04	20	5.55	2.52	14	5.36	2.30
TBLTBSA	45	4.33	1.68	20	5.02	1.92	14	5.75	1.34

An analysis of the DRP mean scores of the students placed into Reading 071, Reading 081 and Reading 019 by the NDRT (see Table 4.9) shows that the students' scores did increase in magnitude on the DRP (41.9, 48.8, 49.8) respectively as the difficulty level of the reading class increased. This means that students who scored higher on the DRP tended to be those students who also scored higher on the NDRT to the extent that they had been placed in more difficult developmental reading classes.

An analysis of the TBLTB-MC and TBLTB-SA mean scores of the students placed into Reading 071, Reading 081 and Reading 091 by the NDRT (see Table 4.10), shows that the students' scores did increase in magnitude on the TBLTB-MC (3.58, 4.80, 5.46 respectively) and the TBLTB-SA (3.50, 4.96, 5.38 respectively) as the difficulty level of the reading class increased. This means that students who scored higher on the TBLTB-MC and TBLTB-SA tended to be those students who also scored higher on the

NDRT to the extent that they had been placed in more difficult developmental reading classes. As expected, students placed in Reading 071, the lowest developmental reading class, scored the lowest on both formats of the TBLTB. Students placed in Reading 081, scored higher than the students in Reading 071. And the students in Reading 091 scored the highest. Still those students who were placed into Reading 091, were only able to correctly answer 6 items of the 10 multiple-choice questions possible and 5 items of the 11 short-answer questions possible on the average. Even though these students were placed in the highest developmental reading class, most of them really did not do that well on the TBLTB. It appears that the Dreams reading passage on the TBLTB was too hard for many of these students to comprehend and that they probably were not ready at that time for university-level course work. It appears that there was a tremendous floor effect with the TBLTB for students placed into the three developmental reading classes by the NDRT.

An analysis of the TBLTB mean scores of the students placed into Reading 071, Reading 081 and Reading 091 by the DRP (see Table 4.11), shows that the students' scores did increase in magnitude on the TBLTB-MC (4.07, 5.55, 5.36 respectively) and on the TBLTB-SA (4.33, 5.02, 5.75 respectively) as the difficulty level of the reading class increased. (There was a variation in the mean scores of students in Reading 081 to Reading 091, showing a decrease of .19, although both scores were greater than the mean scores of students placed in Reading 071). A mean score pattern similar to that with the NDRT was obtained with the DRP. However, the TBLTB mean differences across the

three classes in terms of DRP placement were not nearly as great as those obtained by the NDRT placement. Students placed in Reading 071, the lowest developmental reading class, scored the lowest on the TBLTB. Students placed in Reading 081, scored higher on the TBLTB than the students in Reading 071. And the students in Reading 091 scored the highest on the TBLTB than those in the other two developmental reading classes. Still those Reading 091 students only completed five items correctly on the average for both the TBLTB-MC and the TBLTB-SA. Even though those students had been placed in the highest developmental reading class, they did not really score that well on the TBLTB. It appears that the Dreams reading passage on the TBLTB was too hard for these students to comprehend, and they were not ready for university-level course work at the time they participated in this study. It appears that there was a tremendous floor effect with the TBLTB for students placed into the three developmental reading classes by the DRP.

Analysis of Variance (ANOVAs) were calculated among the scores of the tests according to placement in the three developmental reading courses. The results are found in Table 4.12.

Mean differences were found between the NDRT and the DRP for students placed into three developmental reading classes by the NDRT at the .01 level of confidence.

Differences were found among TBLTB-MC mean scores for students placed into the three developmental reading courses by the NDRT at the .01 level of confidence (see Table 4.13).

Table 4.12

Analysis of Variance among the DRP Mean Scores for Students Placed into Three Developmental Reading Classes by the NDRT

Source of Variation	DF	Squares	Square	F	of F
Between Groups	2	917.977	458.988	4.517	.0137
Within Groups	85	8636.102	101.601		
Total	87	9554.079			

Table 4.13

Analysis of Variance among the TBLTB-MC Mean Scores for Students Placed into Three Developmental Reading Classes by the NDRT

Source of Variation	DF	Sum of Squares	Mean Square	F	Sig. of F
Between Groups	2	36.219	18.109	3.5008	.0348
Within Groups	81	419.018	5.173		
Total	83	455.238			

Differences were found among TBLTB-SA mean scores for students placed into the three developmental reading courses by the NDRT at the .01 level of confidence (see Table 4.14).

Table 4.14

Analysis of Variance among the TBLTB-SA Mean Scores for Students Placed into Three Developmental Reading Classes by the NDRT

Source of Variation	DF	Sum of Squares	Mean Square	F	Sig. of F
Between Groups	2	38.710	19.355	7.348	.0012
Within Groups	81	213.349	2.633		
Total	83	252.059			

Table 4.15 shows differences among TBLTB-MC mean scores for students placed into the three developmental reading courses by the DRP at the .05 level of confidence.

Differences were found among TBLTB-SA mean scores for students placed into three developmental reading courses by the DRP at the .05 level of confidence (see Table 4.16).

Table 4.15

**Analysis of Variance among the TBLTB-MC Mean Scores for Students Placed into
Three Developmental Reading Classes by the DRP**

Source of Variation	DF	Sum of Squares	Mean Square	F	Sig. of F
Between Groups	3	49.473	16.491	3.251	.0260
Within Groups	80	405.764	5.072		
Total	83	455.128			

Table 4.16

**Analysis of Variance among the TBLTB-SA Mean Scores for Students Placed into
Three Developmental Reading Classes by the DRP**

Source of Variation	DF	Sum of Squares	Mean Square	F	Sig. of F
Between Groups	3	25.247	8.415	2.968	.0368
Within Groups	80	226.812	2.835		
Total	83	252.059			

The ANOVAs demonstrated significant mean differences among students placed into the three SWCC developmental reading courses: (1) the NDRT and the DRP; (2) the NDRT and the TBLTB-MC and (3) the NDRT and the TBLTB-SA at the .01 level of confidence. Two other combinations of tests showed significance at the .05 level: (1) the DRP and the TBLTB- MC, and (2) the DRP and the TBLTB-SA. In summary, the ANOVA analyses of the above two test combinations mean that the test score averages for the students in Reading 071 differed significantly from those scores obtained by students placed into Reading 081 and Reading 091. The average scores on a test increased significantly as the course difficulty increased. Since the students had been placed into those three developmental reading courses by either the NDRT or the DRP, it suggested that the two tests did assess literacy competencies in a similar manner that the TBLTB assessed literacy competencies.

Qualitative Data

In addition to the quantitative research done in this study, three students were interviewed to add depth to the study. Providing rich descriptions of three of the subjects should give the reader a clearer picture of the students who made up the sample.

Three selected students were individually interviewed toward the end of the Spring 1998 semester, one semester after they had participated in the study during the Fall 1997 semester. The three students were orally interviewed and were asked to tell about their perceptions of themselves as a student, and how they perceived their own literacy skills. They told the researcher about other courses they had taken during the Fall 1997

and Spring 1998 semesters. They also described how they thought they had performed in their courses.

The three students selected for this in-depth analysis were chosen because they exhibited different score patterns on the three reading assessments investigated in this study. The first student was chosen because she scored low on the NDRT and was advised to take a developmental reading course, but scored high on the other two assessment instruments. The second student was selected because he scored low on the NDRT and the TBLTB, but scored higher on the DRP. The third student was chosen because she scored low on all three assessment instruments and had been placed in the lowest reading class. (The names used in this section are not the names of the students interviewed.)

Ann

The first student, Ann, was 22 years of age and had been out of college for a year when she participated in the study. She had attended a business school right out of high school and had earned a certificate in Hospitality Management. After working a few years in that career, she found herself dissatisfied and stressed with the job. She had thought, on occasion, about becoming a bilingual teacher and when a friend who was teaching in a local school asked her to volunteer in her 5th grade classroom, Ann agreed. She realized she really enjoyed being in the classroom with children, so she decided to go to college and work toward a teaching degree in Bilingual Education, since she was bilingual.

When Ann arrived at the college to apply and register for classes, she was surprised to find out that she had to take some placement tests before she could register. With some reluctance, she went to the testing center, because as she reported, she really hates tests. She did poorly on the NDRT and was told she needed to take a developmental reading course because of her low score. When we talked about this, she said she was upset at first. She said, “I know how to read.”

However, Ann felt there were many reasons why she did poorly on the NDRT test. In the first place, she didn't know that she would have to undergo testing before she could register for classes, and that made her nervous. Then she found out that the test was timed, and she was even more stressed by that. One of her major strategies for understanding and remembering text is rereading, an effective skill. But timed tests allow little extra time for rereading. The NDRT is strictly timed and students very rarely complete the 100 vocabulary items in 15 minutes, or the reading passages in the 20 minutes. This was the case with Ann. She completed 33 items correctly on the vocabulary subtest, which is the number of items completed nationally by students on this subtest of the NDRT. The students in this study on an average completed 22 items. Ann completed 15 items of the 36 items on the comprehension subtest.

Ann's total raw score on the NDRT was 63 (33 on the vocabulary and 30 [2 points per question] on the comprehension), indicating that her reading skills ranged from 7.8 to 9.9 grade level, and she was placed in Reading 081, the second level of developmental reading classes. However, during the first week of school, she earned a raw score of 59

on the DRP that indicated that she was reading very near the college level. If that assessment instrument had been used initially for placement purposes, Ann would have taken Reading 091, the third level of developmental reading classes instead of Reading 081.

However, when Ann was tested on the TBLTB, she was one of the few students who passed both sections of that assessment. Ann correctly answered nine of the eleven multiple-choice questions and earned 75% on the short-answer section. Reading over the short-answer section of the test, it was noted that Ann responded to each question in a complete sentence and demonstrated that she understood the college-level passage on Dreams. The results of the TBLTB suggested that Ann was reading at the college level and that possibly the NDRT had placed her at the wrong reading level.

Another indication that Ann was evaluated and placed inappropriately by the NDRT, was that she was taking two college-level courses, writing and sociology during that Fall semester, and she completed them both with an A grade, earning a 4.0 GPA for the semester.

The following semester, Ann decided not to take another reading course and instead enrolled in 14 hours of college-level courses, including writing, history, biology and psychology. She again earned a 4.0 GPA.

It seems that Ann was not evaluated appropriately by the NDRT and was placed too low according to that assessment instrument. The NDRT indicated that she was an underprepared reader, and would probably have trouble in college-level courses. This

proved to be in error and forced Ann to take a class that she really did not need. The DRP said she was reading near the college-level but still needed some help with developing literacy skills. Again, this evaluation was wrong. The TBLTB was the only instrument that showed that Ann was able to read and comprehend a college text. The TBLTB also showed that Ann had relatively good writing ability. It appears that Ann's low score on the NDRT was probably due to a combination of anxiety and the nature of the test's form. In summary, there was no agreement among all the tests in either evaluation or placement.

Guillermo

The second student, Guillermo, was a 23 year old male from Mexico, who had studied in the United States educational system for only five years. When he took the NDRT, he scored a 31 for a total raw score, indicating that his English literacy skills were considerably below the college level, maybe as low as the fifth grade level. Guillermo was placed in the lowest reading course, Reading 071. When Guillermo was asked about being placed in that class, he said he agreed with it. Even though he said he was "a very good reader in Spanish," he had never read an entire book in English on his own. However, later on, he scored a 52 on the DRP, which indicated that he was not reading at the college level, but that he should be placed in a higher Reading class, Reading 081.

In summary, the NDRT and the DRP tests did agree that Guillermo did not have college reading skills, but they were not consistent in their developmental reading course placement recommendations. Interestingly, Guillermo did very well in the Fall 1997

reading class and the teacher recommended that he skip the next developmental reading class and go to the third level. So even though the NDRT indicated that Guillermo needed a developmental reading class, maybe Reading 081 might have been a more appropriate class, the class that the DRP indicated as the appropriate class for Guillermo.

Or one could look at Guillermo through the work of Vygotsky (1962) who contended that each student has two levels of development: the actual level of development and the potential level or the zone of proximal development. The standardized tests could only evaluate the actual level of Guillermo's literacy development. They could not measure his potential development. Only in the classroom, in the social environment, as Guillermo interacted with the teacher and other students, was he able to demonstrate the wealth of literacy background knowledge that he had. Through the aid of the teacher and other students, Guillermo was able to share that knowledge verbally and in writing.

When Guillermo took the TBLTB, he failed to demonstrate college-level reading skills on the Dreams passage from the psychology text and sufficiently answer the multiple-choice comprehension questions. Guillermo answered two of the multiple-choice questions correctly and earned a 45% on the short-answer section. Guillermo did not even attempt to answer two of the short-answer questions. The questions that he did answer correctly were written in complete sentences. Three questions that he attempted to answer, but failed to answer correctly, were short phrases or words that Guillermo just

lifted from the text. It was apparent that he did not comprehend a large amount of the reading passage.

What was interesting, though, Guillermo did better on the short-answer questions than on the multiple-choice questions. Guillermo was able to discuss in writing his understanding of the text better than pick out the right answers on the multiple-choice questions. This difference could be explained by the fact that he completed the multiple-choice items first, and then responded to the short-answer questions, making possible improvement due to a *gaining effect*. It was the second time he had responded to questions about the passage, and maybe working through the multiple-choice section helped Guillermo clarify the meaning of the passage.

But another fact should also be considered. Guillermo took a college writing course each semester and completed both of them successfully. His writing abilities were probably higher than his reading and comprehension abilities. This is in line with the research of Wallace (1986), who noted that second language learners do not learn a second language in the same way as their first language. When children learn their first language they first speak the language, then read it, then finally write the language. However, the order is often different in the learning of a second language. One may learn to read and write a second language first before ever speaking it. This might be one reason why Guillermo was able to respond to the short-answer questions more effectively than the multiple-choice questions. The first two tests, the NDRT and the DRP, didn't even offer this option to Guillermo. Possibly these two tests were inappropriate

assessment instruments for second language learners, such as Guillermo, who may need different ways of communicating their knowledge.

More importantly, other factors are coming into play in this scenario as more is learned about the social environment of Guillermo. that no paper-and-pencil, timed test, is designed to tap. Even though research suggests that students must be academically prepared for college to be successful, that is not the only factor that determines success. Another factor that came into play with Guillermo was motivation. Guillermo has dropped in and out of college several times over the last few years. He has had to overcome many detrimental family and social situations, but he is determined to get an education. He has decided that he wants an education, and he has a goal to become a bilingual sports broadcaster. As Guillermo said, "If I have to stay up all night and study and work harder than everyone else, I will finish college. I will succeed." Motivation and commitment should get Guillermo through college as Suarez-Orozco (1995) has found in immigrant students. At least it appears that it has helped Guillermo successfully get through the first two semesters at SWCC.

Liliana

The third student, Liliana, was the youngest of the three interviewees, only 19 years of age, and right out of high school during the Fall 1997 semester. She was still living at home with her parents and her eight brothers and sisters at the time. Even though English is her second language, she was born in the United States and had attended school all her life in the United States. She reported that she had not

participated in Bilingual or ESL classes in high school. Her English proficiency appeared to be that of a native speaker.

Liliana scored 38 points on the NDRT and was advised to take the Reading 071, the lowest level of developmental reading classes, indicating that she was reading below the fifth grade level. Her raw score of 44 on the DRP also placed her in the Reading 071 class. In addition, she scored very low on the TBLTB, completing two multiple-choice items and four short-answer questions correctly. Liliana answered all of the short-answer questions in complete sentences. She answered two of the short-answer questions totally right, and four others only partially right. What was apparent in her written answers was that she was able to respond to the literal questions by just lifting the answers straight out of the text. However, she was unable to answer questions that required her to summarize or synthesis material from several paragraphs. In conclusion, all three tests were in agreement in their evaluation of Liliana's literacy skills and both the NDRT and the DRP placed her in the same reading class, Reading 071, which appeared to be appropriate.

At the beginning of the semester, Liliana shared that she thought she was not a very good reader and that in reality she really hated to read. She said that she only read things that she was forced to read in high school. She looked upon reading as pronouncing the words, and she said that reading aloud was very stressful to her. Although she participated in small collaborative groups, she rarely spoke to the whole group, until near the end of the first semester. She requested that she not be made to stand up in front of the entire class and share her book that she had read during the

semester. This was a surprise because she seemed to be very self-assured and easy going. But from the interview and from observing her in the classroom throughout the semester, it was learned that she was quite unsure of herself and her abilities when literacy was involved.

Liliana shared an experience from grade school. The class had divided up into teams for a baseball spelling bee. When it came time for her turn to spell a word, she was “so nervous” and her “hands were sweaty.” Her teammates just stared at her and they said, “Here’s another out.” Of course, she didn’t spell the word correctly. “I felt so stupid and dumb. I let my team down.” She felt those kinds of put-downs really had something to do with her not thinking she was a good reader or speller.

Maybe this class was good for Liliana. It was a small group and the class did a lot of cooperative group activities. In the small groups, Liliana could practice being a leader. During the interview, she shared how much she enjoyed working in the groups and thought she had “learned a lot.” More importantly, she was becoming a reader as the class progressed. She not only was reading books on her own, but she admitted, during the interview, that she was going to book stores and had bought several books. She had never done that before and she was obviously proud when she told me about the books that she had bought. In addition, she was practicing reading aloud to her younger brothers and sisters. Although she still found herself getting nervous, even reading aloud to her family, she was happy that she was reading better and faster. As research indicates, the best way to become a better reader, is to read (Maxwell, 1997; Smith, 1988).

Liliana earned an A in the class, but that was the only course she was taking that semester. The following semester, she not only took a reading course but also a writing course and a math course. When she was asked her how she was doing in her courses, she said “fine, and especially math, because I’m really good at math.” Even though she had done well on all the tests in her math class, she probably would not get an A grade in that course, because she admitted she was a “procrastinator” and “did not do her homework regularly.” And she generally “studied for tests only the night before the test.” When I asked her why she didn’t do her homework, she said she worked 20-30 hours per week, and it was almost impossible to do homework at home. There were so many people in the family, and she shared a room with a younger sister. When I asked her if there was another way to get her homework done, she said, “Yeah, I could really stay after class and study in the library.” But she admitted that she would always change her mind once class was over and she would just go home. She realized when she started taking more courses that were harder, she would have to make a commitment to doing homework.

Liliana did not seem as motivated and goal oriented as the other two students at the time of the study. She was younger and really had not decided what she wanted to do in terms of a career. During both the Fall and Spring semesters of that academic year she reported that she was going to go talk to an advisor and maybe take some of the interest inventory assessments in the career center. However, near the end of the spring semester she had still not done that.

Liliana is still young and that could account for her indecision about school and not working as hard as possible in college. Unlike Guillermo, she was living at home in a supportive environment. She told me even though she helped out at home with the kids and the cleaning and the cooking, her parents continually told her that school came first. Research suggests that family support plays a big part in students being successful in college and that could be the factor that keeps Liliana in school and passing her courses.

It is also possible that the developmental reading course was just what Liliana needed, a supportive class that helped her not only improve her reading but to gain some confidence and ease into college. A regular college course probably would have been too hard for Liliana and maybe she would have failed her courses and quit college. By being placed into the appropriate reading class, Liliana was able to become a more confident reader and student and seemed to be excited about continuing in college.

CHAPTER 5

SUMMARY, CONCLUSIONS, IMPLICATIONS and RECOMMENDATIONS

Summary

The purpose of the study was to investigate the assessment instruments that a southwestern community college (SWCC) used to evaluate its incoming students' academic literacy skills and to determine how well the results of those tests placed students in either college-level courses or developmental courses. Three college-level academic literacy instruments were investigated: the Nelson-Denny Reading Test (NDRT), the Degrees of Reading Power (DRP) and the Turrentine/Bradley Literacy Testing Battery (TBLTB). Three scores of the NDRT were investigated: the NDRT Vocabulary Subtest (NDRT-V), the NDRT Comprehension Subtest (NDRT-C), and the NDRT Total Test (NDRT-T). In addition, two response formats of the TBLTB were studied: the TBLTB Multiple-Choice Format (TBLTB-MC) and the TBLTB Short-Answer Format (TBLTB-SA). This study investigated the interrelationships of the NDRT, the DRP and the TBLTB for reliability, concurrent validity and extent of agreement for placement into developmental reading courses. Specifically, the purposes for this research study were to determine:

1. The correlations among the NDRT, the DRP and the TBLTB.
 2. The extent of agreement among the NDRT, the DRP and the TBLTB in terms of placement into SWCC's developmental reading courses.
-

The literature was reviewed concerning the types of students that are currently enrolled in community colleges. In general, a change was found in the demographics for the typical college student of the 1990s. The modern community college student is now academically, culturally, linguistically and socially more diverse than the student of a few decades ago.

Assessing incoming students' basic academic literacy skills was found to be a common practice in higher education. Once community colleges assess and identify underprepared students, then appropriate placement into courses that match their academic preparation is a key strategy for improving and increasing student success. That is why finding reliable and valid tests to assess basic academic literacy skills was found to be the first step in the community college assessment/placement process. However, the literature indicated that the reading assessment instruments that are now widely used in higher education and in particular by SWCC, are problematic, especially when used in isolation. However, excellent ways to improve the assessment process for incoming college students were highlighted in the literature.

Data were collected in the Fall 1997 semester at SWCC. The academic literacy skills of 88 students were assessed with the NDRT, the DRP and the TBLTB and those three tests were investigated for reliability, concurrent validity and extent of agreement for placement into the three developmental reading courses, Reading 071, Reading 081, Reading 091 and Reading 112 a college-level reading course, that SWCC had established at the time of the study.

Pearson Product-Moment correlations were calculated among the three tests and significant positive linear correlations were found between all of the tests at the .01 level. However, a number of the tests' intercorrelations were found to be too low for practical validity or reliability purposes. The relatively low obtained correlations could be explained by (1) the relatively few tests items on the TBLTB and (2) a floor effect. In the case of the TBLTB, it was apparent that the test was too hard for most of the students because virtually all of them failed the test. Most of the scores were at the lower end of the continuum, demonstrating little variation. When variability is low, relatively low correlations usually are the result.

The Kappas obtained for pairings of the three tests demonstrated that two of the tests showed concurrent placement validity: the NDRT with the TBLTB-MC and the TBLTB-MC with the TBLTB-SA. The extent of agreement among pairings of these tests exceeded .01 level of confidence. The other tests did not show sufficient extent of agreement for reading developmental course placement purposes. However, the extent of agreement was only statistically significant, not significant in a practical sense. For example, only 34 of the 88 students or 39% of the sample were placed in the same developmental reading class by both the NDRT and the DRP. Sixty-one percent of the students were placed differently by these two tests. Since both the NDRT and the DRP are widely used by community colleges, and since they disagreed regarding the course placement of over half of the students in the sample, the question of which of the two tests would be best to use remains unanswered.

In the case of the agreement between the NDRT and the TBLTB-MC regarding placement of students into the three developmental reading classes, again the statistical significance was at the .01 level of confidence. However, of the 85 students evaluated by both the NDRT and the TBLTB-MC, there was agreement in the placement of 33 students or 39% of the sample. Sixty-one percent of the students were placed differently by these two tests.

A Kappa analysis demonstrated that the extent of agreement between the TBLTB-MC and the TBLTB-SA was great enough to reach the .01 level of confidence. However, there was agreement in the placement of only 29 students or 35% of the sample. The remaining 55 students or 65% of the sample were placed in different classes by the two tests.

The TBLTB was chosen as a criterion measure to evaluate incoming students' basic literacy skills. It was assumed that a reading passage taken from an actual college-level text might be more appropriate for assessing college student's academic literacy skills. Also, including both multiple-choice and short-answer items matched the types of activities that college students would be required to do in a college-level course. Using this criterion assessment instrument provided a means of comparing agreement in the placement of students into college level reading classes with the NDRT and the DRP.

It is assumed, however, that institutions would hope to place a greater percentage (35-39%) of students in the appropriate classes, whether they be developmental or regular college-level courses. The consequences of either overplacing or underplacing students

could be detrimental to some students. If students are placed in a class that was too advanced, they might become so discouraged that they would drop out. On the other hand, students who are underplaced might also drop out because they were not being challenged. Again, accurate and appropriate assessment and placement is vital to the admissions process of incoming students to higher education.

The ANOVAS demonstrated significant mean differences among students placed into the three SWCC developmental reading classes by : (1) the NDRT and the DRP, (2) the NDRT and the TBLTB-MC, and (3) the NDRT and the TBLTB-SA at the .01 level of confidence. Two other combinations of tests showed significance at the .05 level of confidence: (1) the DRP and the TBLTB-MC, and (2) the DRP and the TBLTB-SA.

Mean scores were computed for both the TBLTB-MC and the TBLTB-SA according to: (1) the placement of students by the NDRT into the three developmental reading classes, and (2) the placement of students by the DRP in the three developmental reading classes. The mean scores of the students placed into Reading 071, Reading 081 and Reading 091 by the NDRT increased as the difficulty level of the reading class increased. In a similar manner, the mean scores of the students placed into Reading 071, Reading 081 and Reading 091 by the DRP also increased as the difficulty level of the reading class increased. This means that students who scored higher on the NDRT and the DRP scored higher on the TBLTB. However, we cannot say that if a student was assessed and placed by one of the tests, that the outcome would be repeated by another of the tests. Combining the results of both the Kappas and the ANOVAs, the tests did not

have extent of agreement for placement purposes into the three developmental reading courses.

Conclusions

Four conclusions can be drawn from the results of this research study: (1) all three tests, the NDRT, the DRP and the TBLTB, did agree that the sample of students did not possess college-level reading skills; (2) there was little agreement among all three of the tests regarding the placement of students into the three developmental reading courses; (3) the TBLTB consisting of both a multiple-choice and short-answer section provided more information about the students' reading and writing abilities than either the NDRT or the DRP; and (4) the qualitative research demonstrated that even though some students were not prepared academically for college-level course work, they still succeeded in college.

For a large number of students, the three assessment instruments often yielded different results, and placed these students into different levels of developmental reading courses. However, the majority of the SWCC students in the sample, as demonstrated by the scores on all three assessment instruments, were found to be reading below college-level. All three of the tests agreed in that respect, that these students were not prepared for college-level reading experiences. Thus, the students' performance on the TBLTB validated, to some extent, the NDRT and the DRP.

The TBLTB-SA provided more specific information about student's individual reading abilities than either the NDRT and the DRP. An analysis of the short-answer

responses revealed not only students' reading comprehension abilities but also students' abilities to respond in writing to questions about a reading passage. Written responses provided information with regard to sentence structure, spelling, grammar and depth of understanding. It was apparent that some students possessed only literal understanding of parts of the reading passage as words were just lifted straight from the text to respond to the short-answer questions. At times, those words made no sense, demonstrating the student did not even understand the question. In contrast, other students understood underlying factors and were able to write complete, coherent sentences that demonstrated they were able to compare and contrast issues, summarize, draw conclusions and synthesize information. It was concluded that the TBLTB provided more information about students' reading comprehension skills and writing abilities than either the NDRT or the DRP. The writing section of the TBLTB provided diagnostic information that identified specific strengths and weaknesses of students' abilities that could be used to guide instruction.

The qualitative research demonstrated that some students, even though underprepared, still succeed in college. Both Ann and Guillermo scored below college-level reading on both the NDRT and the DRP and had been placed into developmental reading classes. However, they were simultaneously enrolled in other college-level courses. At the end of the semester they both received passing grades in all of their classes. The interviews informed the researcher that both of these students were highly motivated to succeed in college, and were willing to do whatever was necessary to pass

their classes. This information is supported by research that confirms that other factors besides cognitive factors come into play when considering success in college.

Implications

The three tests, the NDRT, the DRP and the TBLTB did agree that most of the students in the sample did not possess college-level reading skills. However, they did not agree on the placement of students into the three developmental reading courses. This is not too surprising. The NDRT, the DRP and the TBLTB were not designed to place students into three distinct developmental reading classes. They were designed to determine general reading abilities of college students. It appears that colleges should only use tests for what they are designed to do, rather than using them for a variety of purposes.

It appears that a criterion-referenced test consisting of both a multiple-choice and short-answer section would be a better assessment instrument to use than a norm-referenced assessment instrument when assessing incoming college students' academic literacy skills. Criterion-referenced instruments provide more information about the reading and writing abilities of students and determine if students are prepared for college-level activities, since the test represents a true reading assignment that college students might have to do in a freshman college class.

From the data gathered, it appears that it would be unwise for a community college to use only one test to assess incoming students' academic literacy skills for the purpose of making decisions about placing students into developmental reading courses.

Other factors, aside from cognitive factors, play a role in students being successful in college.

The case study analysis revealed that success or failure in college appeared to be determined by a number of factors, some of which were related to academic literacy skills, and some of which were not. The three students who were interviewed scored quite differently on the three assessment instruments. Ann, in particular, scored low on the NDRT, because of test anxiety and was underplaced. In reality, Ann was prepared for college-level course work, but because SWCC used only one test score for placement purposes, she was required to take an unnecessary developmental reading course. This inappropriate placement could have caused Ann to become discouraged with college and drop out. Yet, she was motivated to get a college degree so she could become a bilingual teacher. Personal motivation may have been the factor that kept Ann in school.

If SWCC had used more than one assessment instrument for placement purposes, most likely Guillermo would have been placed into a higher developmental reading class. He too was motivated to complete college and begin his career as a sports broadcaster. And even though his academic literacy skills were found to be below college-level, he was succeeding in college at the time of the study. Being underprepared alone was not found to always be the determining factor for success in college.

All three of the assessment instruments assessed and placed Liliana into the lowest developmental reading class. This was the only instance of the three students

interviewed where the three tests demonstrated reliability and concurrent validity in terms of developmental reading course placement.

Recommendations

More research investigating the reliability and validity of basic academic literacy assessment instruments needs to be done. As student diversity continues to grow in higher education, it is important to determine if current assessment instruments that colleges and universities are using to assess their incoming students are appropriate for these culturally and linguistically diverse students.

A criterion-referenced assessment instrument, such as the TBLTB, might be more appropriate for assessing incoming students' basic academic literacy skills. More importantly, this instrument required students to respond to short-answer questions in addition to the multiple-choice questions during the assessment process. Writing is a vital part of the college curriculum in most content areas, and instructors need to know the writing abilities of their students. This portion of the test gave more information about the reading and writing skills students possessed and the ones they lacked. However, it was apparent, in the case of the TBLTB, that more test items were needed on the test to increase its reliability and validity, plus offer more variability in test results.

In addition, community colleges need to use more than one test score in making academic decisions for incoming students. Not only academic abilities influence college students' success. Factors, such as age, motivation, family support may affect a student's performance in college. That is why it is important for advisors and instructors to be

aware of the strengths and weaknesses of incoming students so students can be advised appropriately to either enroll in regular college-level course work or to provide them with an intervention, such as developmental courses or tutoring. That is why a comprehensive assessment/placement program is recommended. A comprehensive program would include a variety of services for the students, plus the use of a number of factors to make placement decisions. For example, if a student is a recent high school graduate, in addition to test scores and high school GPA, advisors could look at the courses that a student had taken in high school. If the student had taken more than the minimum requirements for high school, including a number of science and math courses, then it might be determined that the student was prepared for the rigors of college course work.

If the student is a returning student who has been out of high school for some time, or has never completed high school, then other factors might be considered in conjunction with the test scores. Motivation, goals, family responsibilities and work commitments all play a role in the success or failure of college. Finding out about these affective factors might give insight into the needs of a student.

More and more community colleges are moving towards a more comprehensive assessment/placement program. Community colleges have found that comprehensive programs play a major role in assisting students in being successful in college and ultimately lowering drop out rates.

APPENDIX A
PARTICIPANT'S CONSENT FORM



DISSERTATION RESEARCH STUDY

Researcher: Mary D. Shelor
The University of Arizona

Participant's Consent Form

I am being asked to read the following material to ensure that I am informed of the nature of this research study and of how I will participate in it, if I consent to do so. Signing this form will indicate that I have been so informed and that I give my consent. Federal regulations require written informed consent prior to participation in this research study so that I can know the nature and the risks of my participation and can decide to participate or not to participate in a free and informed manner.

Purpose:

I am being invited to voluntarily participate in this research project. The purpose of the study is to investigate the agreement among college reading instruments and their relation to developmental course placement. The study will also describe the student who attends Desert Vista Campus of Pima Community College and takes developmental reading courses.

Procedure:

If I agree to participate, the researcher will have access to my reading scores on the Nelson Denny, the Degrees of Reading Power, the Turrentine/Bradley Informal Reading Survey and any other assessments that might be administered during the Fall, 1997 semester. In addition, a variety of surveys will be completed focusing on demographics, enrollment status, reading interests and literacy background. My participation is strictly voluntary, and I may withdraw consent at any time.

Confidentiality:

All test scores and survey information will be kept strictly confidential. Personal names will not be used.

I have read and understand the above information. My signature below indicates that I agree to participate in this research study.

Name (Please print)

Date

Signature

APPENDIX B
DREAMS READING PASSAGE
AND
TEST QUESTIONS

Dreams as Mental Experiences

One way of looking at dreams is as conscious mental experiences. According to an old account that goes back to the Greek philosopher Aristotle, the dream happenings are mental re-evocations of sights and sounds that occurred during the dreamer's waking life. Aristotle believed that the succession of these dream images from the past is experienced as real while it occurs because during sleep there is no competition from the clamor of waking reality and because the intellect is "dulled" during sleep (Aristotle, ca. 330 B.C.).

Later investigators tried to relate what people dream about to what happens to them both before and during sleep. One question concerns the effect of recent waking experiences. Aristotle was apparently correct in his belief that such recent events often re-emerge in dreams. This is especially so when the recent waking experience was highly emotional. For example, soldiers who have just gone through intense battle stress may relive their combat terrors in nightmare dreams (Pai, 1946).

Some writers have suggested that the dream images from the past are supplemented by external events that impinge upon the sleeper in the present. A widely cited example is the alarm clock, which is often said to turn into a peal of church bells or a fire engine in the dream. To test this hypothesis, several investigators have studied the effects of applying various forms of external stimulation during sleep. Numerous sleepers have been shaken, tickled, splashed with water, and shouted at-all to discover whether they would later report a dream that referred to these experiences. Sometimes they did. An example of a dream reported on awakening after an experimenter shouted "Help" into the sleeper's ears: "I was driving along the highway at home. Heard yelling and we stopped. A car was turned sideways in the road. I went down and saw the car was turned over on the side of the road.... There was a woman badly cut. We took her to the hospital" (Hall, 1966, p. 6). It is hard to resist the conclusion that, at some level, the dreamer heard and understood the shout "Help" even while asleep and then incorporated it within his dream narrative.

Dreams as Behavior

Dreams as conscious, mental experiences are essentially private, they go on "inside" the individual. As such, dreams can be regarded as a form of behavior that is looked at from within, as if the actor were observing his own actions. Yet psychologists study most aspects of behavior from "outside," for much of what we do is directly apparent and overt and can therefore easily be seen by others. Humans and animals act. They run and fly and scurry about; they eat and fight and mate; they often perform new acts to attain their ends.

Overt Behavior

Can we study dreaming from the outside by taking this action-oriented view? On the face of it, the prospects don't seem too bright, for during sleep the body is by and large immobile. Even so, there is a way. For there is one thing the sleeper does while dreaming that is overt and can be observed from the outside. He moves his eyes.

This fact emerged after it became clear that there are two kinds of sleep: quiet sleep and active sleep. During quiet sleep, both breathing and heart rate are slow and regular, and the eyes are motionless. But during active sleep the pattern is different. Breathing and heart rate accelerate, and-most characteristic of all-the eyes move back and forth behind closed lids in quick irregular darts. Periods

of quiet and active sleep (often called REM sleep because of the Rapid Eye Movements) alternate throughout the night, and with a total of perhaps ninety minutes devoted to REM (Figure 1.7).

The crucial factor about REM sleep is that this is the period during which dreams occur. When subjects—the persons whose behavior is being studied—are aroused during REM sleep, about 85 percent of the awakenings lead to reports of a vivid dream. In contrast, subjects awakened from non-REM sleep recall dreams much less often (Dement, 1974).

There is little doubt that the rapid eye movements go along with dreaming. But some investigators take a further step. They suggest that these movements provide a clue to what the sleeper is dreaming of. In their view, the eye movements indicate that the sleeper is “looking at” whatever he sees in his dream world. Some evidence for this intriguing (though quite controversial) hypothesis comes from studies which show the direction of the eye movements observed during a given REM period is appropriate to what the subject recalls having seen while dreaming. For example, when the predominant direction of eye movement was up and down, one subject dreamed that he had thrown basketballs, looking up at the net and shooting, then looking down to pick another ball off the floor. In contrast, another REM period, in which the eye movements were mostly from side to side, produced a dream in which the subject watched two people throwing tomatoes at each other (Dement and Kleitman, 1957).

Biological Underpinnings of Behavior

Evidence based on eye movements is one clue to the nature of dreaming. Another clue concerns its biological basis. Most psychologists take it for granted that whatever we do or think has some physical basis in the activity of our brains. Some promising first steps in this direction pertain to dreams. These came after the development of a number of methods for monitoring what people do when they are fast asleep. An important tool is the electroencephalogram, or EEG, which measures the overall activity of the sleeper's brain. The EEG traces the quick fluctuations of electrical activity over time, and its graphic record is sometimes called brain waves. Happily for sleep investigators, it soon proved possible to attach the multitude of required electrodes to the subjects without disturbing their sleep (Figure 1.8).

The results showed that the EEG patterns of sleep and waking differ markedly. As the subject falls into deeper and deeper stages of sleep, the brain waves become slow, large, and rather regular, indicating a lower level of brain activity. But this holds only for periods of quiet sleep. As this is interrupted by active (that is REM) sleep, the EEG becomes quite similar to that found when the subject is awake. This makes good sense, for it suggests that during REM sleep the brain is reasonably aroused and active—as well it should be since this is the time when we are busily engaged in dreaming.

Dreams as Cognition

Like many other psychological phenomena, dreams reflect what we know, what we have experienced, remembered, or thought about—activities that psychologists call cognition. To be sure, the dream happenings didn't really take place. We didn't really fly through the air or have tea with Queen Elizabeth. But the components of the dream were surely drawn from the dreamer's own knowledge, which

contains information about flying and the queen of England. How was this knowledge retrieved and woven into the dream story? How was the dream recalled on later awakening? And why is it that most of us remember so few of our dreams?

Some psychologists have tried to attack these and related questions by asking about the factors that make for better dream recall. They have come up with some evidence that people who remember more of their dreams are more likely to have better and sharper visual mental images in their waking life; perhaps their dreams are more memorable because they are experienced in a more vivid pictorial form (Corey et al., 1975). Another factor is the extent to which the dream experience is interfered with by what happens immediately after the sleeper awakes. In one study, subjects were asked to call the weather bureau immediately after waking up; after this, they had to write down a detailed description of any dreams they had that night. The results showed that weather reports and dreams don't mix. The subjects who made the call generally remembered that they had had a dream, but most of them could not remember what it was that they had dreamed about (Cohen and Wolfe, 1973).

Dreams and Social Behavior

Human life is very solitary but is spent among a world of others—strangers and friends, partners and rivals, potential and actual mates. What holds for waking existence, holds for dreams as well. Most of them involve interactions with others. Some feature themes of aggression, such as competition, attack, and submission. Others concern friendship and sometimes sex. But whatever the plot, the cast usually includes some others. More than 95 percent of our dreams are peopled with others, and most revolve around our relations with them (Hall and Van de Castle, 1966).

Dreams and the Culture

Dreams concern major themes in the person's own life, but they take place within a larger framework, the dreamer's own culture. In our society, a common dream is of appearing naked among strangers and being embarrassed. But such a dream would be unlikely among Australian aborigines who wear no clothing. Nor are many urban Americans likely to have nightmares in which they are chased by cows, which happens to be a common dream in western Ghana (Barnoouw, 1963)

Culture affects not only what the dream is about but also how the dreamer thinks about it when she recalls it later on. In some societies, including our own, dreams are generally dismissed as nonsensical fancies, irrelevant to real life. Many preliterate cultures have a different point of view (Figure 1.9). Some regard dreams as supernatural visions and behave accordingly. Others take dreams very seriously even though they think of them as naturally occurring events. The Senoi, a tribe in Malaya, act as if they had all taken several courses in psychoanalysis. They believe that dreams indicate something about their inner lives and can provide clues for heading off problems before they become serious. Every morning Senoi children tell their father what they dreamed about the night before. The father then helps the children interpret their dreams. These may reveal some incipient conflict with others, as in a dream of being attacked by a friend. If so, the father may advise the child on how to correct matters, for example, by giving the friend a present (Stewart, 1951).

Dreams and Internal Conflict

The social aspect of dreaming lies at the heart of a famous (and controversial) theory of dreams proposed by Sigmund Freud. According to Freud, dreams are the product of an elaborate clash between two contending forces—the unconscious primitive urges of our biological heritage and the civilizing constraints imposed by society. In dreams we sometimes see one, sometimes the other side of the battle. Various forbidden impulses—mostly sex and aggression—emerge, but they are soon opposed by the thou-shalt-nots of our early upbringing. The result is a compromise. The forbidden material breaks through but only in a stealthy, censored masquerade. This disguise explains why dreams are so often odd and senseless. Their senselessness is only on the surface, a cunning mask that lets us indulge in the unacceptable wish without realizing that it is unacceptable (Freud, 1900).

According to Freud, some distortions involve various transformations of the unacceptable themes. One is symbolism. For example, he believed that sexual urges often emerge in symbolic guises. Thus, in his view, dreams of riding horses or walking up a staircase often mask erotic wishes. Here the symbols presumably bear some resemblance to what they symbolize. The rhythmic movements of rising and falling in the saddle are similar to those of sexual intercourse, while ascending a staircase may be reminiscent of the way in which sexual passion mounts to a peak (Figure 1.10).

Freud argued that these and many other symbolic transformations are the dreamer's way of smuggling the forbidden wish past the inner censor's eye. He believed that such defenses refer back to early childhood when the parents set up the various prohibitions that still haunt the adult in the present. Seen in this light, dreams reflect important social processes that pertain to the past, to the way in which the major social commandments were installed in each of us by society's first agents, our parents.

Dreams and Human Development

Thus far we have discussed dreams as they are experienced by adults. But of course dreams occur in childhood as well as in adulthood. Psychologists who are concerned with the course of human mental development have considered the different ways in which children and adults think about their dreams.

Developmental psychologists want to know how children acquire the basic intellectual operations that are part of adult human thought—how they learn to count, to understand that events have causes, and so on. For example, they ask how children learn that there is a difference between two realms of phenomena, those that we call subjective (thoughts, beliefs, and of course, dreams) and those that we call objective (the world of tangible things “out there”). To ask how this distinction is made is another way of asking how we attain our adult notion of objective reality, how we come to know that the tree in the garden—unlike a dream—will still be there after we blink our eyes.

The distinction is by no means clear in early childhood. Thus young children initially have great trouble in distinguishing dreams from waking life. A three-year-old awakes and tell her parents how much she loved the elephants at the circus yesterday. The parents correct her: she had not been at the circus yesterday. But the child indignantly sticks to her story and appeals to her brother for collaboration, for "he was there too." When her brother shakes his head in denial, she begins to cry, angrily insisting that she told the truth. Eventually she learns that there is a whole group of experiences that older people call "just dreams," no matter how real they seem to her (Levy, J., 1979).

The fact that the child finally recognizes the circus elephants-and the nightmare robbers and witches-as dreams does not mean that she has acquired an adult conception of what dreams are. Young children tend to think of them as physical objects. When asked whether dreams can be tall, a four-year-old replied, "Yeah. How tall? *Big, big, big* (spreads arms). Where are dreams? *In your bedroom*. In the daytime? *No, they're outside...* What are they like? *They're made of rock*. Could they be heavy? *Yeah, and they can't break either*" (Keil, 1979, pp. 109-10).

It's quite a while before children think of dreams the way adults do. By six or seven, they believe that dreams are sent through the air, perhaps by the wind or by pigeons. Eventually they recognize that, as one eleven-year-old put it, "You dream *with the head* and the dream is *in the head*" (Piaget, 1972).

This realization that dreams are subjective events is no small achievement. As we will see later, the recognition that some experiences are subjective, inner happenings is not limited to dreams, but extends to many other conceptual attainments about the basic nature of the physical and psychological universe.

Dreams and Individual Differences

There is a further aspect of dreams. They are a reflection of the fact that people are different. People vary in what they characteristically do and think and feel. And some of these differences between people are reflected in their dreams. Some simply pertain to the differing circumstances in the dreamers' lives. This point was made some two thousand years ago by the Roman poet Lucretius who noted that at night lawyers plead their cases, generals fight their battles, and sailors wage their wars with the winds (Woods, 1947).

More interesting are differences that reveal something about the personalities of the dreamers. An example is a comparison of the dreams of normal people and of patients with a diagnosis of schizophrenia, a condition generally regarded as the most serious psychiatric disorder of our time. The difference between the two groups was enormous. The schizophrenic reported dreams that were highly bizarre and often morbid. The dreamer is eaten alive by an alligator; there are nuclear wars and world cataclysms. Themes of bodily mutilations were fairly common, as in a dream in which a woman killed her husband and then stuffed parts of his body into a camel's head. In contrast, the normals' dreams were comparatively mild and ordinary. The results fit in with what we know about schizophrenia. Schizophrenics often jump from one idea to the other without maintaining a firm line of thought. As a result, their behavior often appears bizarre. And they often report great personal turmoil and distress. It seems that their extremely bizarre and morbid dreams are simply an exaggeration of a condition already present in their waking life (Carrington, 1972).

Perspectives on Psychology

We have seen that dreams can be looked at as conscious, mental experiences, as overt behaviors, as aspects of cognition, as indications of social patterns, as reflections of human development, and as expressions of the dreamer's individuality. What holds for dreams holds for most other psychological phenomena: They can all be viewed from several perspectives. Each perspective is valid but non is complete without the others, for psychology is a field of many faces and to see it fully, we must look at them all.

Given the many-faceted character of psychology, it is not surprising that those who have contributed to it come from many quarters. Some had the proper title of psychologist with appropriate university appointments in that discipline, including two of its founding fathers, Wilhelm Wundt of Germany and William James of the United States. But psychology was not built by psychologists alone. Far from it. Among its architects are philosophers, beginning with Plato and Aristotle and continuing to our own time. Physicists and physiologists played important roles and still do. Physicians contributed greatly, as did specialists in many other disciplines, including anthropology, and more recently linguistics and computer science. Psychology, the field of many faces, is by its very nature a field of many origins.

NAME: _____

Passage 1-Multiple Choice Questions (Circle the letter preceding the correct response)

1. According to Aristotle, dreams:
 - a. are outlets for creative energies
 - b. are manifestations of physical discomfort during the day
 - c. are of telepathic origin
 - d. are reoccurrences of events in the real world
 - e. a and d

 2. Outside stimuli:
 - a. always find their way into dreams
 - b. never find their way into dreams
 - c. sometimes find their way into dreams
 - d. are all that dreams are made of
 - e. very rarely find their way into dreams

 3. During sleep we can tell that the sleeper is dreaming when we observe that:
 - a. the sleeper's heart rate accelerates
 - b. the sleeper's breathing slows down
 - c. the sleeper's eyes dart back and forth behind closed lids
 - d. a and c
 - e. a and b

 4. Subjects dream during:
 - a. all types of sleep
 - b. quiet sleep
 - c. the end of the night only
 - d. REM sleep
 - e. b and d

 5. Electroencephalograms (EEG) show that:
 - a. the brain waves during periods of dreaming show a pattern that is slow, long, and regular
 - b. the brain waves during periods of dreaming are no different than those seen during non-dreaming periods
 - c. brain wave patterns during dream periods are similar to those seen when the subject is awake
 - d. researchers have been unable to attach the EEG leads to subjects without disturbing their sleep
 - e. b and c
-

Passage 1-Multiple Choice Questions (p. 2)

6. Better dream recall is related to:
- the ability of the subject to have better and sharper recall of visual mental images in their waking life
 - non-interference of other factors immediately after the sleeper awakes.
 - the intelligence of the dreamer
 - a and b
 - a and c
7. Culture:
- has influence only on dream content.
 - determines only how people interpret dreams
 - influences the dream content and interpretation
 - may or may not influence our dreams at all depending on the particular culture.
 - none of the above.
8. In Freud's view, dreams:
- are a result of physical discomfort during sleep
 - represent a compromise between biological urges and social constraints
 - serve as a reminder of events of the day
 - pertain mostly to current conflicts in real life
 - b and d
9. According to Freud, dreams contain:
- symbolic of fulfillment of forbidden wishes
 - prophecies of the future
 - undisguised urges
 - all of the above
 - a and c
10. Young children think of dreams as:
- physical objects
 - a special reality sent from outside
 - purely subjective experiences
 - both a and b
 - none of the above
11. The dreams of schizophrenic individuals are often:
- more sedate than those of others
 - more exciting than those of others
 - more bizarre than those of others
 - b and c
 - not substantially different from those of others

APPENDIX C
PUBLISHER'S PERMISSION LETTER

May 11, 1997

Fred Courtright
W. W. Norton Company
New York, New York

Dear Mr. Courtright,

I am working on my dissertation at the University of Arizona and would like permission to use a six page passage from the text, Psychology, 3rd Edition, by Henry Gleitman. About one hundred students will be asked to read a passage about Dreams and then answer a number of multiple choice and short answer questions. I am comparing the results of this reading assessment to the results of two different standardized reading tests that are used in community colleges for assessing academic literacy skills of incoming students.

Please let me know if this permission can be granted. My mailing address is: 5933 East 20th Street, Tucson, AZ 85711

Thank you for considering my request.

Sincerely,

Mary D. Shelor
Mary D. Shelor

PERMISSION GRANTED for a fee of \$ 0
payable to W. W. Norton & Company, Inc. (Tax
ID 14-5718330). Full credit to our work is
required on each copy reproduced.

Sincerely,

Fred Courtright
Frederick T. Courtright
Rights and Permissions Manager

REFERENCES

- Abraham, A. A., Jr. (1987). A report on college-level remedial/developmental programs in SREB States. Atlanta, GA: Southern Regional Education Board.
- American College Testing Program. (1985). American College Testing and American Association of Community Junior Colleges to cooperate in national surveys of community, junior and technical colleges. Iowa City, IA: American College Testing Programs.
- Astin, A. . (1975). Preventing students from dropping out. San Francisco: Jossey-Bass.
- Astin, A. W. (1991). Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education. New York: American Council on Education and Macmillan.
- Boylan, H. R. (1983). A review of diagnostic reading tests. Reading in Developmental Education, 1, 1-6.
- Boylan, H. R., Bingham, E. L., & Cockman, D. C. (1988). Organizational patterns for developmental education programs. Review of Research in Developmental Education, 5(4), 1-4.
- Boylan, H. R., & Bonham, B. (1992). The impact of developmental education programs. Research in Developmental Education, 9(5), 1-4.
-

Boylan, H. R., Bonham, B., & Bliss, L. (1992). National study of developmental Education: Students, programs and institutions of higher education--Summary report. Boone, NC: National Center for Developmental Education.

Boylan, H. R., Bonham, B., & Bliss, L. (1994). Characteristic components of developmental programs. Research in Developmental Education, 11(1), 1-4.

Bray, D. (1987). Assessment and placement of developmental and high-risk students. In K. M. Abrendt (Ed.), Teaching the developmental student: New directions for community colleges (pp.33-48). San Francisco: Jossey-Bass.

Brown, J. I., Bennet, J. M., & Hanna, G. (1981). Examiner's manual: The Nelson-Denny Reading Test, Forms E & F. Chicago, IL: Riverside Publishing.

Bruning, R. (1985). Degrees of reading power. In J. V. Mitchell (Ed.), The ninth mental measurements yearbook (pp. 443-444). Lincoln, NE: University of Nebraska Press.

Carver, R. P. (1985). Is the degrees of reading power test valid or invalid? Journal of Reading, 28, 34-41.

Chumchall, S. K. (1996). The educational experience of nontraditional age female African American students. In J. L. Higbee & P. L. Dwindell (Eds.), In defining developmental education: Theory, research & pedagogy (pp. 49-62). Cold Stream, IL: National Association for Developmental Education.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, XX, 37-46.

College Board. (1983). Academic preparation for college. NY: The College Board.

Cress, C. (1996). ERIC Review: Measuring success through assessment and testing. Community College Review, 24, 1.

Cummins, R. P. (1981). Test review: The Nelson-Denny Reading Test (Forms E and F). Journal of Reading, 24, 54-59.

Deem, J. M. (1981). The relationship of sociocultural distance to the reading and writing performance of community college basic skills students. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Degrees of reading power, PA and CP Series (1983). New York: The College Board.

Ediger, M. (1991). Reading on the higher education level. (ERIC ED 336 725).

El-Khawas, E. (1989). Campus trends: 1989. Washington, DC: American Council on Education.

Enright, G. (1988, June). Reading and writing test reviews. Paper presented at the California State University, Long Beach, Fourth Annual Institute for Learning Assistance Professionals, Long Beach.

Evangelou, J. (1985). Enrollment in remedial courses jumps at 63 percent of colleges that offer them. Chronical of Higher Education, (February 13), 3.

Farr, R. (1972). Test reviews: Nelson-Denny Reading Test. In W. Blanton, R. Farr, & J. H. Tuinman (Eds.), Reading tests for the secondary grades: A review and evaluation (pp.31-34). Newark, DE: International Reading Association.

Farr, R., & Carey, R.F. (1986). Reading: What can be measured? (2nd Ed.). Newark, DE: International Reading Association.

Flippo, R. F. (1980). The need for comparison studies of college students' reading gains in developmental reading programs using general and specific levels of diagnosis. In M. L. Kamil & A. J. Mae (Eds.), Perspectives on reading research and instruction. Washington, DC: National Reading Conference. (ED 184 061).

Flippo, R. F., & Caverly, D. C. (1991). College reading & study strategy programs. Newark, DE: International Reading Association.

Fullilove, R. E., & Treisman, P. U. (1990). Mathematics achievement among African-American undergraduates at the University of California-Berkeley: An evaluation of the mathematics workshop program. Journal of Negro Education, 59(3), 463-477.

Gleitman, H. (1991). Psychology (3rd ed.). New York: W.W. Norton & Company.

Grable, J. R. (1988). Remedial education in Texas two-year colleges. Journal of Developmental Education, 12(2), 2-5.

Hanna, G. S. (1985). Review of degrees of reading power. In Mitchell, J. V. (Ed.), The ninth mental measurements yearbook (pp. 444-447). Lincoln, NE: University of Nebraska Press.

- Hills, J., Hirsch, T., & Subhiyah, R. (1990). Issues in college placement. Washington, DC: ERIC/American Institutes for Research.
- Ignash, J. M. (1997). Who should provide postsecondary remedial/developmental education? New Directions for Community Colleges, 100, 5-20.
- Johnston, P. (1983). Reading comprehension assessment: A cognitive basis. Newark, DE: International Reading Association.
- Jonides, J., & Rozin, P. (1991). Study guide: Gleitman Psychology (3rd ed.). New York: W. W. Norton & Company.
- Kerlin, S. P., & Britz, P. B. (1994). Assessment and diversity: Outcome and climate measurements. In T. H. Bers & M. L. Mittler (Eds.), New directions for community colleges (pp. 53-60). San Francisco: Jossey-Bass.
- Kerstiens, G. (1986). A testimonial on timed testing: Developmental students and reading comprehension tests. In M. Douglass (Ed.), Fiftieth yearbook of the Claremont Reading Conference (pp.267-268). Claremont, CA: Claremont Graduate School.
- Kerstiens, G. (1990). A slow look at speeded comprehension tests. Review of Research in Developmental Education, 7, 3.
- Kiah, C. G. (1993). A model for assessing critical thinking skills. Paper presented at an Annual Student Assessment Conference of the Virginia Assessment Group and the State Council of Higher Education for Virginia, Richmond, VA. (ERIC Document Reproduction on Service No. ED 367 400).

Kibby, M. W. (1981). The degrees of reading power. Journal of Reading, 24, 416-427.

Kirst, M. (1995). Exam confusion. Crosstalk, 5, 1, 3.

Krotseng, M. V., & Pike, G. R. (1991). Cognitive assessment instruments: availability and utilization. In J. O. Nichols, A practitioner's handbook for institutional effectiveness and student outcome assessment implementation. New York: Agathon Press.

Lederman, M. J., Ryzewic, S. R., & Ribaud, M. (1985). Basic skills of entering college freshmen: A national survey of policies and perceptions. Journal of Developmental Education, 9(11) 10-13.

Levine, A., & Cureton, J. S. (1998). When hope and fear collide. San Francisco: Jossey-Bass Publishers.

Levitz, R., & Noel, L. (1995). The earth-shaking but quiet revolution in retention management. Noel-Levitz occasional paper. Iowa City, IA: USA GROUP/Noel-Levitz.

Lucas, C. J. (1996). Crisis in the academy: Rethinking higher education in America. New York: St. Martin's Press.

MacDouglass, A. (1977). The relationship between reading competency and successful performance in the curriculum. Chula Vista, CA: Southwestern College Research Report.

Malinowski, P. A. (1988). A survey of college reading programs in New York State: Diagnosis, placement and program components. (ERIC Document Reproduction Service No. ED 298 989).

Mason, R. (1994). Trends and issues in college reading improvement programs. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement. (ERIC Document Reproduction Service No. ED 375 384).

Maxwell, M. (1991). Evaluating academic skills programs: A source book. Kensington, MD: M. M. Associates.

Maxwell, M. (Ed.). (1994). From access to success. Clearwater, FL: H & H Publishing Co., Inc.

Maxwell, M. (1997). Improving student learning skills. Clearwater, FL: H & H Publishing Co., Inc.

McMillan, P. K. (1994). Assessment from the state perspective. In T. H. Bers & M. L. Mittler (Eds.), New directions for community colleges (pp. 5-11). San Francisco: Jossey-Bass.

McNabb, I. (1990). Course placement practices of American postsecondary institutions, 1985-86. Unpublished manuscript, The American College Testing Program, Iowa City, IA.

McWhorter, K. T. (1994). Academic reading (2nd ed.). New York: Harper-Collins.

Morante, E. A. (1987). A primer in placement testing. In D. Bray & M. Belcher (Eds.), Issues in student assessment: New directions for community colleges (pp. 55-63). San Francisco: Jossey-Bass.

Morante, E. A. (1989). Selecting tests and placing students. Journal of Developmental Education, 13(2), 2-6.

Morante, E. A. and others. (1982). Report to the Board of Higher Education of results of the New Jersey Basic Skills Placement Testing and recommendations on instruction and curriculum. Trenton: New Jersey State Board of Higher Education. (ED 232 716).

Morante, E. A., Faskow, S., & Menditto, I. N. (1984) The New Jersey basic skills assessment program. Journal of Developmental & Remedial Education, 7(2), 2-4. 32.

National Center for Education Statistics. (1990). Digest of education statistics. (NCES Publication No. 91-660). Washington, DC: U.S. Government Printing Office.

National Commission on Testing and Public Policy. (1990). From gatekeeper to gateway: Transforming testing in America. Boston College, Chestnut Hill, MA: National Commission of Testing and Public Policy.

Nelson-Denny Reading Test. Forms E and F. (1981). Boston, MA: Houghton Mifflin.

Pascarella, E. T., & Terenzini, P. T. (1991). How college affects students: Findings and insights from twenty years of research. San Francisco: Jossey-Bass.

Payne, E. M., & Lyman, B. G. (1992). Issues affecting the definitions of development education. In J. L. Higbee & P. L. Dwinell, (Ed.), Defining developmental education: Theory, research & pedagogy (pp. 11-20). Cold Stream, IL: National Association for Developmental Education.

Perkins, D. (1984). Assessment of the use of the Nelson Denny Test. Forum for Reading, 15(2), 64-67.

Prus, J., & Johnson, R. (1994). A critical review of student assessment outcomes. In T. H. Bers & M. L. Mittler (Eds.), New directions for community colleges (pp. 68-83). San Francisco: Jossey-Bass.

Raygor, A. L. (1978). The Nelson-Denny Reading Test (Forms C and D). In O.K. Buros (Ed.), The eighth mental measurements yearbook (Vol. 2) (pp. 1209-1211). Highland Park, NJ: The Gryphon Press.

Rosenblatt, L. M. (1978). The reader, the text, the poem: The transactional theory of literary work. Carbondale, IL: Southern Illinois University Press.

Ross, E. P., & Roe, B. D. (1986). The case for basic skills programs in higher education. Fastback 233. Bloomington, IN: Phi Delta Kappa. (ED 273 166).

Roueche, J. E. (1984). Between a rock and a hard place: Meeting adult literacy needs. Community and Junior College Journal, 54(7), 21-24.

Roueche, J. E., & Baker, G.A. (1987). Access & excellence: The open-door college. Washington, DC: The Community College Press.

- Roueche, J. E., & Roueche, S. D. (1993). Between a rock and a hard place. Washington, DC: The American Association of Community Colleges.
- Roueche, J. E., & Snow, J. (1977). Overcoming learning problems: A guide to developmental education in college. San Francisco: Jossey-Bass.
- Rounds, J. C. (1984). Assessment and placement in language arts courses: three major tests. Journal of College Reading and Learning, XVII, 94-100.
- Sammons, R., & Davey, B. (1994). Assessing students' skills in using textbooks: The textbook awareness and performance profile (TAPP). Journal of Reading, 37(4) 280-286.
- Sawyer, R. (1990). Placement testing. Unpublished manuscript. Boone, NC: Appalachian State University.
- Seybert, J. A. (1994). Assessment from a national perspective: Where are we, really? In T. H. Bers & M. L. Mittler (Eds.), New directions for community colleges (pp. 23-30). San Francisco: Jossey-Bass.
- Skinner, E. F., & Carter, S. (1987). A second chance for Texans: Remedial education in two-year colleges. Tempe, AZ: Arizona State University.
- Smith, F. (1988). Understanding reading: A psycholinguistic analysis (4th ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Smith, S. P., & Jackson, J. H. (1985). Assessing reading/learning skills with written retellings. The Reading Teacher, 28(7), 622-630.

- Sternberg, R. J. (1991). Are we reading too much into reading comprehension? Journal of Reading, 34(7), 539-544.
- Stoker, H. (1987). In D. J. Keyser & R. C. Sweetland (Eds.), Test critiques (Vol. VI). Kansas City, MO: Test Corporation of America.
- Suarez-Orozco, C. (1995). Transformations: Immigration, family life and motivation among Latino adolescents. Stanford, CA: Stanford University Press.
- SWCC Handbook. (1996). 1995/96 fact book (Volume 24), Institutional Research Office.
- Tennessee Board of Regents. (1991). The effectiveness of the Tennessee Board of Regents' academic assessment, placement, and remediation program: A follow-up report on the 1986 cohort of remedial/developmental students. Nashville, TN: Author.
- Tierney, R. J. (1985). Review of the Nelson-Denny Reading Test, Forms E and F. In J. V. Mitchell, Jr. (Ed), The ninth mental measurements yearbook (pp.1035-1039). Lincoln, NE: University of Nebraska Press.
- Tinto, V. (1987). Leaving college: Rethinking the causes and cures of student attrition. Chicago: University of Chicago Press.
- Tinto, V. (1988). States of student departure: Reflections on the longitudinal character on student leaving. Journal of Higher Education. 59, 438-455.
- Touchstone Applied Science Associates. (1995). DRP handbook, G & H test forms. Brewster, NY: TASA.
-

Turrentine, P. A., & Bradley, J. M. (1997). The Turrentine/Bradley battery for evaluating the literacy competencies for undergraduate college students. Unpublished manuscript. Tucson, AZ: University of Arizona.

U. S. Department of Education. (1996). National Center for Educational Statistics: Condition of education, 1996. (NCES 96-304). Washington, DC: U. S. Government Printing Office.

Van, B. (1992). Developmental education: Its past and future. Journal of Learning Improvement, 1(2), 21-28.

Van Blerkom, M. L., & Van Blerkom, D. L. (1981). Misleading findings from the Nelson-Denny Reading Test. Journal of Reading, 24, 736-737.

Vygotsky, L. S. (1986). Thought and language. Cambridge, MA: The MIT Press.

Wallace, C. (1986). Learning to read in a multicultural society: The context of second language literacy. New York: Pergamon Institute of English.

Wood, N. V. (1989). Codifying literacy: Identifying and measuring reading competencies in statewide basic skills assessment programs. Journal of College Reading and Learning, 22, 1.

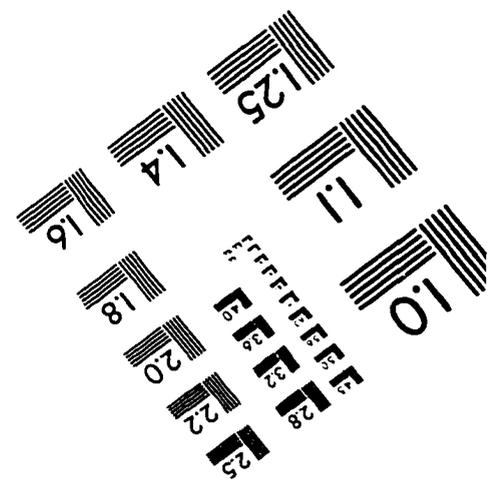
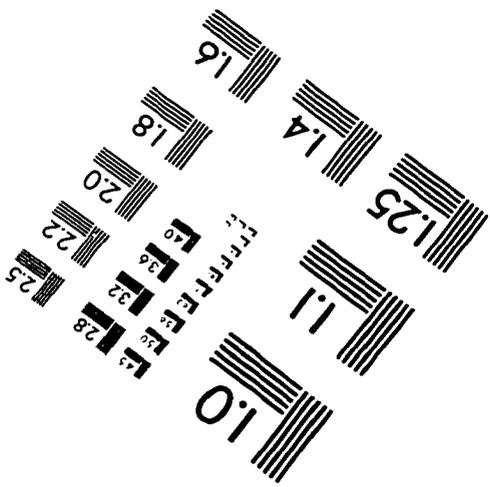
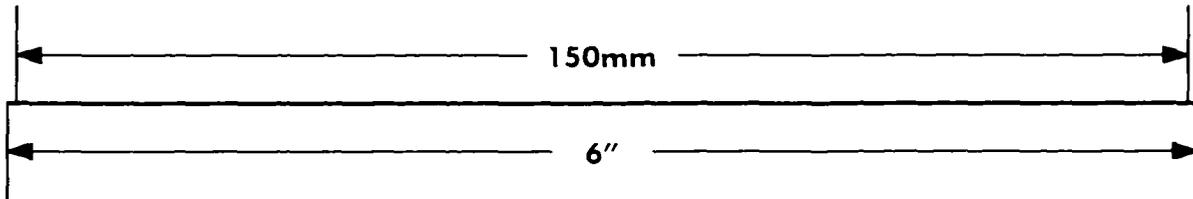
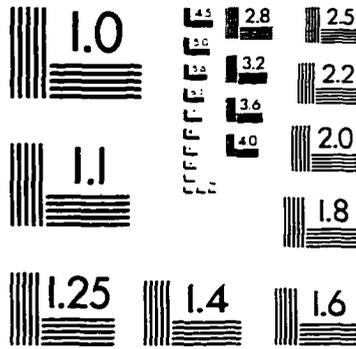
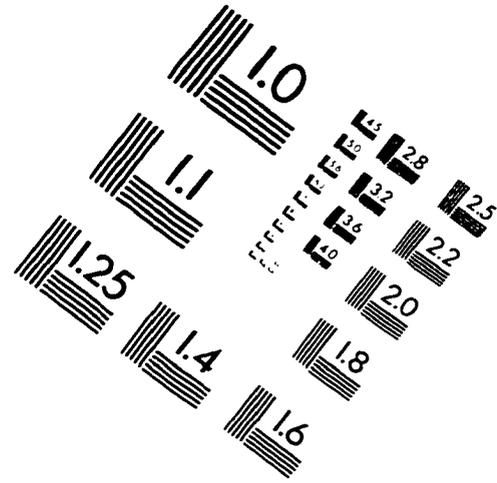
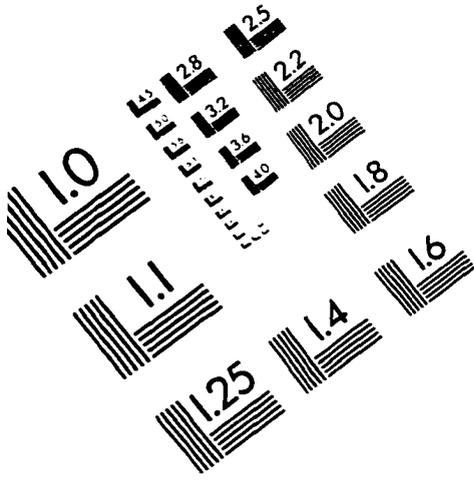
Wood, N. V. (1989). Reading tests and reading assessment. Journal of Developmental Education, 13(2), 14-19.

Wood, P. H., & Bennett, T. (1990). The results of six years of testing college students with Form E of the Nelson-Denny Reading Test. (ED 322487).

Wyatt, M. (1992). The past, present, and future needs for college reading courses in the United States. Journal of Reading, 36(1), 10-20

Ysseldyke, J. E. (1985). Review of the Nelson-Denny Reading Test, Forms E and F. In J. V. Mitchell, Jr. (Ed.), The ninth mental measurements yearbook (Vol. 2). Lincoln, NE: The University of Nebraska Press.

IMAGE EVALUATION TEST TARGET (QA-3)



APPLIED IMAGE, Inc
 1653 East Main Street
 Rochester, NY 14609 USA
 Phone: 716/482-0300
 Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved