# Florida International University FIU Digital Commons

FIU Electronic Theses and Dissertations

University Graduate School

11-8-2012

Examining the Validity of the GRE General Test Scores and Undergraduate GPA for Predicting Success in Graduate School at a Large Racially and Ethnically Diverse Public University in Southeast Florida

Myung Sook Hyun Florida International University, mhyun001@fiu.edu

**DOI:** 10.25148/etd.FI12121002 Follow this and additional works at: https://digitalcommons.fiu.edu/etd

#### **Recommended** Citation

Hyun, Myung Sook, "Examining the Validity of the GRE General Test Scores and Undergraduate GPA for Predicting Success in Graduate School at a Large Racially and Ethnically Diverse Public University in Southeast Florida" (2012). *FIU Electronic Theses and Dissertations*. 797.

https://digitalcommons.fiu.edu/etd/797

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fu.edu.

# FLORIDA INTERNATIONAL UNIVERSITY

## Miami, Florida

# EXAMINING THE VALIDITY OF THE GRE GENERAL TEST SCORES AND UNDERGRADUATE GPA FOR PREDICTING SUCCESS IN GRADUATE SCHOOL AT A LARGE RACIALLY AND ETHNICALLY DIVERSE PUBLIC UNIVERSITY IN SOUTHEAST FLORIDA

A dissertation submitted in partial fulfillment of the

requirements for the degree of

## DOCTOR OF EDUCATION

in

# ADULT EDUCATION

#### AND

## HUMAN RESOURCE DEVELOPMENT

by

Myung Sook Hyun

To: Dean Delia C. Garcia College of Education

This dissertation, written by Myung Sook Hyun, and entitled Examining the Validity of GRE General Test Scores and Undergraduate GPA for Predicting Success in Graduate School at A Large Racially and Ethnically Diverse Public University in Southeast Florida, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

Benjamin Baez

Paulette Johnson

Tonette S. Rocco

Thomas G. Reio Jr., Major Professor

Date of Defense: November 8, 2012

The dissertation of Myung Sook Hyun is approved.

Dean Delia C. Garcia College of Education

Dean Lakshmi N. Reddi University Graduate School

Florida International University, 2012

© Copyright 2012 by Myung Sook Hyun

All rights reserved

# DEDICATION

This dissertation study is dedicated to all my spiritual teachers who gave me the strength to finish this project.

#### ACKNOWLEDGMENTS

I would like to thank all the individuals who made this research study possible. First of all, I would like to thank each of my dissertation committee members, Professors Dr. Benjamin Baez, Dr. Paulette Johnson, Dr. Thomas G. Reio, Jr., my dissertation chair, and Dr. Tonette S. Rocco, for their contributions to my research study. Without their support and guidance, I would not have been able to complete this research study.

Dr. Baez, I am very grateful to you for your caring support. At every turn, you encouraged me in my dissertation study with advice, and kind words. At our first meeting, you expressed a great interest in my study and gladly agreed to serve on my dissertation committee. You recommended two books for reference which were very helpful in broadening my perspective on my research. You also helped me to resolve an administrative issue involving my dissertation proposal defense, and helped me move forward with my study. You responded to all my emails and requests promptly even during a semester break, all of which facilitated my progress.

Dr. Johnson, Statistical Consultant at Florida International University, and the professor for my statistics courses, I would like to extend my gratitude to you for generously sharing your expertise and dedicating more than a hundred hours of your time to help me to organize and analyze the most complicated data for this research study. Thanks to you, the data analysis was completed, and the study yielded useful new knowledge to our school and to GRE validation studies. You can be assured that findings from this study will serve our graduate school administrators as a reference in reviewing or updating educational policies and procedures. Evidence related to differential predictions from this study will be of interest to many future researchers. I hope you are

v

proud of your contribution to this research study. I felt privileged working with, and learning from you during the entire process of data analysis.

Dr. Reio, my dissertation chair, I am most grateful to you. During each step of my dissertation study, you provided me with guidance, insights, and instructions that helped me to improve my dissertation study. You gave me directions when I was lost; you encouraged me to persevere when I felt stuck. There are just so many things you have done for me to get this dissertation study finished.

Dr. Tonette S. Rocco, I would like to express my sincere appreciation to you for many things you have done to help me complete my dissertation study. First, you allowed me to sit in your dissertation proposal writing course so that I could resume my dissertation study. I benefitted greatly from the instructions and guidance provided by your course. In addition, I was inspired to continue my study by many students in your dissertation writing group who demonstrated that with your guidance we could finish our dissertation study if we worked diligently. It was quite a learning experience, especially in the early stages of this dissertation study, and I felt privileged to be a part of this group.

Finally, I would like to extend my special thanks to Drs. Jeff Gonzalez and Hiselgis Perez at the Office of Planning and Institutional Research (OPIR), for providing me with the student data sets that made this research study possible. I am also thankful to Rasheed Oladele, a computer programmer, who prepared the data sets and answered my questions.

### ABSTRACT OF THE DISSERTATION

# EXAMINING THE VALIDITY OF THE GRE GENERAL TEST SCORES AND UNDERGRADUATE GPA FOR PREDICTING SUCCESS IN GRADUATE SCHOOL AT A LARGE RACIALLY AND ETHNICALLY DIVERSE PUBLIC UNIVERSITY IN SOUTHEAST FLORIDA

by

Myung Sook Hyun

Florida International University, 2012

Miami, Florida

Professor Thomas G. Reio, Jr., Major Professor

The purpose of the study was to determine the degree of relationships among GRE scores, undergraduate GPA (UGPA), and success in graduate school, as measured by first year graduate GPA (FGPA), cumulative graduate GPA, and degree attainment status. A second aim of the study was to determine whether the relationships between the composite predictor (GRE scores and UGPA) and the three success measures differed by race/ethnicity and sex.

A total of 7,367 graduate student records (masters, 5,990; doctoral: 1,377) from 2000 to 2010 were used to evaluate the relationships among GRE scores, UGPA and the three success measures. Pearson's correlation, multiple linear and logistic regression, and hierarchical multiple linear and logistic regression analyses were performed to answer the research questions.

The results of the correlational analyses differed by degree level. For master's students, the ETS proposed prediction that GRE scores are valid predictors of first year

graduate GPA was supported by the findings from the present study; however, for doctoral students, the proposed prediction was only partially supported.

Regression and correlational analyses indicated that UGPA was the variable that consistently predicted all three success measures for both degree levels. The hierarchical multiple linear and logistic regression analyses indicated that at master's degree level, White students with higher GRE Quantitative Reasoning Test scores were more likely to attain a degree than Asian Americans, while International students with higher UGPA were more likely to attain a degree than White students. The relationships between the three predictors and the three success measures were not significantly different between men and women for either degree level.

Findings have implications both for practice and research. They will provide graduate school administrators with institution-specific validity data for UGPA and the GRE scores, which can be referenced in making admission decisions, while they will provide empirical and professionally defensible evidence to support the current practice of using UGPA and GRE scores for admission considerations. In addition, new evidence relating to differential predictions will be useful as a resource reference for future GRE validation researchers.

viii

# TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	
	Background to the Study	2
	Problem Statement	7
	Purpose of the Study	9
	Research Questions	10
	Conceptual Framework	11
	Operational Definitions of Terms	15
	Significance of the Study	18
	Assumptions	20
	Delimitations	21
	Summary	21
П	LITERATURE REVIEW	
	Conceptual Review	
	Rationale for the Prediction	
	Test Theory	
	What Is a Test?	
	What Is a Test Score?	
	How to interpret a test score?	
	Theory of Performance and Knowledge Acquisition	
	Empirical Review	40
	Variables: Predictors and Criterion (Success) Measures	40
	Outline for the Empirical Review	
	Criteria Used in Selecting the Related Literature	45
	Organization of the Selected Literature	46
	Validity Evidence Based on White Student Majority	47
	Summary of the Validity Data Based on White Student Majority .	62
	Validity Evidence Based on Minority Students	64
	Summary of the Validity Data Based on Minority Students	68
	Present Study	69
	Summary	70
III.	METHOD	72
	Research Design	72
	Population	74
	Sample	74
	Variables and Indicators	75
	Predictor Variables	76
	Criterion Variables	77
	Subgroup Variables	78
	Data Collection	

	Data Analysis	79
	Research Question One	79
	Research Question Two	80
	Research Question Three	80
	Research Question Four	81
	Summary	83
IV.	RESULTS	84
	Sample	85
	Demographic Variables	86
	Research Variables	87
	Research Question One: Correlational Analyses	90
	Research Question Two: Correlational Analyses	94
	Research Question Three: Multiple Linear and Logistic	
	Regression Analyses	96
	Research Question Four: Hierarchical Multiple Linear and Logistic	
	Regression Analyses	104
	Summary	115
V.	DISCUSSION	119
	Summary of the Study	119
	Discussion of the Results	122
	Research Question One	122
	Research Question Two	126
	Research Question Three	127
	Research Question Four	129
	Implications for Theory, Research, and Practice	132
	Implications for Theory	133
	Implications for Research	134
	Implications for Practice	136
	Limitations of the Study	139
	Conclusion	140
	REFERENCES	142
	APPENDICES	152
	VITA	171

# LIST OF TABLES

TABLE PAGE
1. Summary of Demographic Variable
<ol> <li>Means, Standard Deviations and Intercorrelations among Variables for Master's Students for Research Question 1</li></ol>
<ol> <li>Means, Standard Deviations and Intercorrelations among Variables for Doctoral Students for Research Question 1</li></ol>
4. Regression Analysis of Master's Students' FGPA on the Composite Predictor for Research Question 3
<ol> <li>Regression Analysis of Doctoral Students' FGPA on the Composite Predictor for Research Question 3</li></ol>
6. Regression Analysis of Master's Students' CGPA on the Composite Predictor for Research Question 3
7. Regression Analysis of Doctoral Students' CGPA on the Composite Predictor for Research Question 3100
8. Logistic Regression of Master's Students' Degree Attainment Status within a 6-year Time Limit on the Composite Predictor for Research Question 3101
9. Logistic Regression of Doctoral students' Degree Attainment Status within a 9-year Time Limit on the Composite Predictor for Research Question 3103
<ol> <li>Hierarchical Multiple Regression Analysis of Doctoral Students' FGPA on the Composite Predictor for Research Question 4109</li> </ol>
<ol> <li>Hierarchical Multiple Logistic Regression Analysis of Master's Students' Degree Attainment Status within a 6-year Time Limit on the Composite Predictor for Research Question 4</li></ol>
12. Descriptive Statistics for Master's Students' GRE Scores and UGPA by Sex and Race/Ethnicity
13. Descriptive Statistics for Master's Students' FGPA, CGPA, and Degree Attainment by Sex and Race/Ethnicity
14. Descriptive Statistics for Doctoral Students' GRE Scores and UGPA by Sex and Race/Ethnicity

15.	Descriptive Statistics for Doctoral Students' FGPA, CGPA, and Degree Attainment by Sex and Race/Ethnicity	.156
16.	Means, Standard Deviations and Zero-Order Correlation Coefficients of GRE Scores and UGPA with FGPA by Sex and Race/Ethnicity	.157
17.	Means, Standard Deviations and Zero-Order Correlation Coefficients of GRE Scores and UGPA with CGPA by Sex and Race/Ethnicity	.158
18.	Means, Standard Deviations and Zero-Order Correlation Coefficients of GRE Scores and UGPA with Degree Attainment within a Given Time Limit by Sex and Race/Ethnicity	.159
19.	Summary of Relationships between GREV, GREQ, UGPA, and FGPA, CGPA, Degree Attainment Status for Research Question 1, 2, and 3 (Master's Degree Level)	.160
20.	Summary of Relationships between GREV, GREQ, UGPA, and FGPA, CGPA, Degree Attainment Status for Research Question 1, 2, and 3 (Doctoral Degree Level)	.161
21.	Summary of Interactions of Race/Ethnicity by the Composite Predictor on Graduate Students' FGPA, CGPA, and Degree Attainment Status for Research Question 4	.162
22.	Validity Data of the GRE Scores and UGPA for Success in Graduate School from Meta-Analytic Studies (White Majority)	.163
23.	Validity Data of the GRE Scores and UGPA with Success in Graduate School by Department from Meta-Analytic Studies (White Majority)	.165
24.	Regression Coefficients of the GRE Scores and UGPA for Success in Graduate School by Department from Meta-Analytic Studies (White Majority)	.166
25.	Validity Data of the GRE Scores and UGPA for Success in Graduate School (African American Majority)	.167
26.	Validity Data of the GRE Scores for Success in Graduate School (American Indian)	.168
27.	Validity Data of the GRE Scores and UGPA for Success in Graduate School (Hispanic and Anglo Students)	.169

28.	Validity Data of the GRE Scores and UGPA for Success in Graduate	
	School (Caucasian Men and Women)	170

### **CHAPTER I**

### **INTRODUCTION**

Benefits of access to graduate education are wide-ranging. For the individual, there are notable financial returns as well as personal and intellectual benefits. For society as a whole, well-educated and highly trained professionals contribute to economic and technological development. It is important therefore, to understand the potential barriers that prevent access to and persistence in graduate education. (NCES, 2007, p. iii)

Most graduate schools require all applicants to submit Graduate Record Examination (GRE) test scores for admission considerations. However, admission tests have long been viewed as a major barrier to higher education especially for minority students (Zwick, 2002). Minority students refer to American Indian/Alaska Native, African American, Asian, and Hispanic (NCES, 2007). While graduate schools may require as much information as they deem necessary to make selection decisions, it is important that problems or concerns that are related to high-stakes decisions such as admission to graduate school be addressed and resolved.

The present study investigated the predictive relationships of the GRE to success in graduate school based on a racially and ethnically diverse graduate student population and its subpopulations (by race/ethnicity, and sex) using regression and correlation methods. This chapter begins with the background to the problem, and problem statement, followed by purpose of the study, research questions, a conceptual framework, operational definitions of terms, significance of the study, assumptions, delimitations, and ends with a summary of the chapter.

#### **Background to the Problem**

Educational Testing Service (ETS) recommends the use of the GRE General Test scores for the selection of applicants for admission to graduate school and for fellowship awards (ETS, 2009). The rationale for such recommendation is that "GRE General Test scores are valid predictors of success in the first year of graduate school for all students" (ETS, 2009, p. 7). This statement implies that those who received higher GRE scores are more likely to be successful in the first year of graduate school than those who did not. But, for this rationale to be justified, two prerequisites must be met. They are the reliability and validity of the test (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1999; Crocker & Algina, 1986). Only when these two requirements are satisfied can such a prediction (or an inference) be made from the test scores. While both qualities (reliability and validity) are required of the test, the first requirement of a test is reliability (Pressley & McCormick, 1995).

#### Reliability

Reliability can be defined as the degree of consistency in test scores of the same test (Crocker & Algina, 1986; Mehrens & Lehmann, 1987), and higher reliability of a test represents higher degree of consistency in test scores. A test score (an observed test score) is composed of a true score and an error. (observed score = true score + error, Gulliksen, 1967). The true score, *T*, or the actual ability of an examinee, is the portion of the observed score that is relatively stable, or unchanging in a set of tests. On the other hand, error, *E*, is "the discrepancy between an examinee's observed test score and his or

her true score" (Crocker & Algina, 1986, p. 110). Error is the portion of the observed score that varies in the set of tests (Gulliksen, 1967; Mehrens & Lehmann, 1987).

For a test score to reveal an examinee's true score, subsequent test scores on the same test should be consistent. If an examinee's test scores are not similar on the same test, the test is not reliable and is of little use. Because the purpose of a test is to use the test scores to differentiate among people and make decisions (Mehrens & Lehmann, 1987), it is important that a test yields consistent results. In other words, a test has to be reliable.

"Depending on which error sources are of greatest concern" (Brown, 1983, p. 67), there are many different ways to calculate many different types of reliability estimates (e.g., measures of stability, measures of equivalence, measures of equivalence and stability, measures of internal consistency). At any rate, reliability is theoretically defined as the proportion of the true score variance in observed score variance (Mehrens & Lehmann, 1987). More specifically, it is the ratio of the variance of a group of individuals' true scores and the variance of the group of individuals' observed scores. The reliability is computed based on variances of a group of individuals' scores.

Reliability is commonly expressed in terms of a reliability coefficient, and it tells us "how much of the variance in observed scores represents variance in true scores" (Brown, 1983, p. 75). For example, if the reliability coefficient of a test is .92, we can say that 92% of the variability in observed scores is due to differences in true scores, and 8% is due to errors of measurement (Brown, 1983, p. 88). On the other hand, if a reliability coefficient is 1.00, the test would be perfectly reliable. Although perfect reliability of a

test would be ideal, "no test is perfectly reliable" (Gay, 1996, p. 145). Thus, perfect reliability is never attained in practice (Brown, 1983).

The reliability coefficients (the reliability indices) of GRE Verbal, and Quantitative reasoning measures were .93, and .92, respectively (The reliability for the Verbal and Quantitative measures of the General Test was calculated using simulated data based on the mean of 12 recent computer-based pools, ETS, 2009). The reliability data on both measures can be considered reliable because "all reliabilities in the .90s are acceptable" (Gay, 1996, p. 300). Therefore, as far as test reliability is concerned, the two GRE General Subtests meet the first requirement.

#### Validity

Another quality required of a test is validity. Validity refers to "the degree to which accumulated evidence and theory support specific interpretations of test scores entailed by proposed uses of a test" (AERA, APA, & NCME, 1999, p. 184). This means, to examine the validity of GRE scores, three tasks need to be taken in order. First, the specific interpretation of test scores which "may take the form of a description, a prediction, or a recommended decision" (Cronbach, 1988, p. 3) by the test publisher need to be clarified as the validity of the test will be defined by the correspondence between the proposed interpretation and the accumulated evidence and theory. For the GRE scores, the specific interpretations entailed by proposed uses of the test takes the form of a prediction. That is, the "GRE General Test scores are valid predictors of success in the first year of graduate school for all students" (ETS, 2009, p. 7). Thus, this prediction has to be clearly understood (e.g., the concept of success). Second, evidence and theory that support the prediction need to be gathered. Last, the degree of the relationship between

the GRE scores and the evidence and theory that supports the prediction needs to be determined.

However, "if the proposed interpretation claims that applicants with higher scores on the test can be expected to exhibit better performance in some activity, it would certainly be reasonable to check on this prediction" (Kane, 2006, p. 18). This is essentially the same as what ETS proposed about the GRE scores. As was discussed before, ETS stated that the GRE General Test scores predict first year academic performance. This implies that those who received higher GRE scores can be expected to exhibit better performance in the first year of graduate school than those who did not. Therefore, it would be reasonable that the present study checks on the proposed prediction.

Checking on the proposed prediction by ETS involves three principal tasks: One is clarifying the interpretations of test scores. The other is gathering relevant data (predictor: GRE General Test scores; criterion: success in the first year of graduate school). Still the other is determining the degree of the relationship between the predictor and the criterion. This means that the operational measure for the criterion, success in the first year of graduate school, is identified, the data are collected from a sample, and the degree of the relationship is determined.

To determine the degree of the relationship, regression and correlation methods have traditionally been used in prediction studies (Ackoff, 1965; Cohen, Cohen, West, & Aiken, 2003; Ezekiel & Fox, 1959; Stevens, 1996). In this method, the degree of a relationship is expressed as a validity coefficient and a regression coefficient. The validity coefficient is a correlation coefficient (Anastasi & Urbina, 1997; Brown, 1983)

and expressed in a number that ranges from -1 to +1 (Gay, 1996). It is denoted by symbol, r, for a single predictor, and R, for multiple predictors. The coefficients of size .10, .30, and .50, regardless of sign, are judged as small, medium, and large correlations, respectively (Cohen, 1988). In any event, for a test to have validity, the coefficient has to be positive (r>0 or R>0) and statistically significant.

A regression coefficient is another type of index that is used to explain the degree (or existence) of relationship. It is denoted by symbol *B* (regression coefficient) and  $\beta$ (standardized regression coefficient). The difference between *B* and  $\beta$  is that while the *B* is computed based on raw scores of the data, the  $\beta$  is computed based on the standardized scores, i.e., the raw scores are "transformed to z scores with a mean of 0 and a standard deviation of 1" (Green & Salkind, 2008, p. 280).

A regression coefficient is a slope weight in a regression line. That is, "the number of units by which *Y* [e.g., first year graduate GPA] changes when *X* [e.g., GREV score] changes by one unit in the linear regression equation" (Ackoff, 1965, p. 211). For example, if the *B* involving first year graduate GPA (*Y*) and GREV (*X*) is .08 (positive) and is statistically significant, it means that the first year graduate GPA (*Y*) changes by .08 units (increase) when GREV changes one unit. However, if the *B* =0, then, it would mean that the first year graduate GPA is not associated with the GREV. In this case, one can conclude that changes in the first year graduate GPA are not associated with changes in the GREV. In short, for the test to have validity, it requires that the regression coefficients be positive (*B*>0), and the relationship be statistically significant.

On the other hand, the same test can have many validities. As was described before, because the data are collected from a sample and the characteristics of a sample

(data or the test scores) generally differ, validity of the same test may vary depending on situations and locations (Brown, 1983; Cronbach, 1971).

#### **Problem Statement**

Numerous predictive validity studies on GRE scores have been conducted by many graduate schools, individual researchers, and the test publisher, ETS, since the first use of the test in 1937 (Zwick, 2002, p. 19). Extensive validity data (validity coefficients) on the GRE scores are available from several meta-analyses (Burton & Wang, 2005; Goldberg & Alliger, 1992; Kuncel, Hazlet, & Ones, 2001; Schneider & Briel, 1990; Thacker & William, 1974; Willingham, 1974; Wilson, 1979). Most meta-analytic studies found that the GRE has positive and statistically significant relationships with a criterion (as measured by first year graduate GPA and the cumulative graduate GPA), although magnitude of the correlations was small. The correlations from these studies fall into the range from .15 to .29 for GREV, and from .15 to .28 for GREQ.

However, some researchers argue about the integrity of the predictive validity of GRE scores (Ingram, 1983; Morrison & Morrison, 1995) while others regard GRE as the major barrier to graduate education for minority students (Sack, 2001). Morrison and Morrison (1995) stated that the test had "such little magnitude [of relationship with graduate grade point average/graduate school success] that it appears that they are virtually useless from a prediction standpoint" (p. 311). Another researcher, Ingram (1983), who examined 10 studies that were conducted over the 1961-1978 time period, concluded that "the GRE might predict success or it might not - we just do not know" (p. 713). The sample sizes of the 10 studies ranged from 31 to 91 with a mean sample size of 57, and the success measures were first year GPA, graduate GPA, faculty ratings, peer

ratings of success, and publications. The validity coefficients varied greatly, ranging from -.32 to .40 for GREV and -.24 to .32 for GREQ.

Researchers who studied the predictive relationships of the GRE scores based on minority students and a population that includes a large number of minority students found puzzling or inconsistent findings: Some researchers found that graduate GPA was not related to GRE scores while others found conflicting results. Whitworth and Barrientos (1990) whose sample (N = 952) included a large number of Hispanic students (N = 320) found that graduate GPA was not significantly related to any of the GRE scores for Hispanic students. In addition, they found that the regression coefficients for the GREV, GREQ, and GRE analytical (GREA) were not significantly different from 0 for Hispanics, indicating that changes in graduate GPA do not tend to be associated with changes in any of the GRE scores. Other researchers, Sampson and Boyer (2001), whose sample (N = 160) included a majority of African American students (N = 144) found that first year graduate GPA was significantly related only to GREV scores, r (160) =.39, (p < .01), but not to GREQ.

It seems that such findings as these may have led some people to question the usefulness of the GRE scores or view the test score requirement in admission considerations as the major barrier to graduate education for minority students. That is, from minority students' perspective, the requirement of the GRE scores can be viewed as something irrelevant and a barrier to graduate education. But, these studies are the two largest published GRE studies (N > 100) based on minority students (Note that 100 is the minimum requirement for a validity study, Anastasi & Urbina, 1997; Cronbach, 1971). In

fact, the validity data based on minority students are extremely limited mainly due to the fact that samples of minority students have been very small (ETS, 2009, p. 7).

The problem is, due to the inconsistent findings and scarcity of available data, it is very difficult to draw any firm conclusions about the predictive validity of the GRE scores for minority students and a population that includes a large number of minority students. We just cannot give the definitive answer to the question of whether the GRE scores predict success in graduate school for these special populations.

The composition of graduate student population has been changing over the past 30 years: Graduate enrollments of minority students have increased steadily, and the trend is projected to continue at least for another decade (NCES, 2008). If GRE scores are to be used in graduate admission considerations and other high-stakes decisions (e.g., fellowship awards) for all students, more research is needed to resolve the issue relative to the predictive validity of the GRE scores for minority students and the population that includes a large number of minority students.

#### **Purpose of the Study**

The purpose of the present study was twofold. One was to determine the degree of relationships between the GRE General Test scores (GRE Verbal Reasoning [GREV], GRE Quantitative Reasoning [GREQ] Test scores), Undergraduate GPA (UGPA), and success measures in graduate school, as measured by first year graduate GPA (FGPA), cumulative graduate GPA (CGPA), and degree attainment status, based on a population that includes a large number of minority students. The other was to determine whether the relationships between GREV, GREQ, and UGPA, as a composite predictor, and success in graduate school, as measured by FGPA, CGPA, and degree attainment status,

differ by race/ethnicity and sex, more specifically, whether the relationships differ between White students and minority students (African Americans, Asians, and Hispanics; and International students. Note that International students were included as one category of a racial/ethnic group in the study), and between men and women.

In this study, the predictors included GREV, GREQ, and UGPA. The criterion was success in graduate school, as measured by FGPA, CGPA, and degree attainment status.

### **Research Questions**

The research questions that guided the study were as follows:

- To what degree do GRE scores taken singly (individual GRE Verbal Reasoning Test score and GRE Quantitative Reasoning Test score), and in total, sum of GREV and GREQ (GRET), predict success in graduate school, as measured by FGPA, CGPA, and degree attainment status?
  - a. To what degree does GREV predict success in graduate school?
  - b. To what degree does GREQ predict success in graduate school?
  - c. To what degree does GRET predict success in graduate school?
- 2. To what degree does UGPA predict success in graduate school, as measured FGPA, CGPA, and degree attainment status?
- 3. To what degree do GREV, GREQ, and UGPA, as a composite predictor, predict success in graduate school, as measured by FGPA, CGPA, and degree attainment status?

4. Do the relationships between GREV, GREQ and UGPA, as a composite predictor, and success in graduate school, as measured by FGPA, CGPA, and degree attainment status, differ by race/ethnicity and sex?

### **Conceptual Framework**

#### **Criterion Model**

The conceptual framework for the present study was based on criterion model (Kane, 2006). Under criterion model, the validity is defined in term of the correlation between the predictor (GREV, GREQ, UGPA) and the criterion of future scores (FGPA, CGPA, and degree attainment status). If the correlation is high, the test is considered to have good predictive validity (Gay, 1996). Pearson's correlation, multiple linear and logistic regression, hierarchical multiple linear and logistic regression methods were used to determine the relationships between the three predictors and the three success measures.

# Pearson's Correlation, Multiple Linear (and Logistic) Regression, and Hierarchical multiple Linear (and Logistic) Regression Methods

Use of Pearson's correlation, multiple linear (logistic) regression, and hierarchical multiple linear (logistic) regression methods provided advantages to the present study: First, the resulting correlation coefficients from a correlation analysis were comparable to those of the past research studies as most previous studies reported the correlation coefficients.

Second, the resulting  $R^2$  (effect size) from a multiple linear (logistic) regression analysis informed us what percentage of the variability in the criterion, FGPA, CGPA, and (degree attainment) were predicted by knowing the GREV, GREQ, and UGPA (Thompson, 2002). Third, the models constructed using hierarchical linear regression were useful in examining the differences in the relationship of the composite predictor (GREV, GREQ, and UGPA) to FGPA and CGPA by race/ethnicity and sex. Fourth, the models constructed using hierarchical logistic regression were useful in examining the differences in the relationship of the composite predictor (GREV, GREQ, and UGPA) to degree attainment status by race/ethnicity and sex.

#### **Factors Affecting the Validity**

Validity indices (validity coefficients and regression coefficients) are computed based on the data (both predictor and criterion scores) collected from a sample. However, several factors involved in the data affect the degree of the test-criterion relationship (validity). They include the sample, restriction in range of scores, and the characteristics of the criterion variables (Breland, 1979; Brown, 1983; Kuncel et al., 2001).

The sample. General requirements for selecting a good sample include that samples are selected randomly, are "representative of the target population so as not to introduce any further biases, and large enough to allow for stable and statistically significant data" (Brown, 1983, p. 122).

However, samples used in most predictive validity studies on admission tests are usually preselected. That is, the data (the scores) used in these studies are collected on the group of students who are already selected, and enrolled in graduate programs, eliminating those of the unselected applicants (probably the lower range of scores). This pre-selection sampling method reduces the magnitude of the validity (Dawes, 1975) because "the magnitude of a correlation depends on the range of individual differences that is, the variance of the predictor and criterion scores" (Brown, 1983, p. 104). What it

means is, if the range of scores is restricted (if the scores of a group do not vary significantly from each other), the variance of the scores will be small, and this will reduce the magnitude of the correlation. In other words, eliminating the lower range of the scores (those of the rejected applicants) and using the selected data will limit the variance of the predictor and criterion scores (scores of the admitted students are generally not so different from each other). And, the restricted range of the scores will "attenuate the magnitude of the validity coefficients" (Hartnett & Willingham, 1980, p. 282).

This unique characteristic of the sample in the predictive validity study is one of the factors that reduce the magnitude of the test-criterion relationship (e.g., GRE scores – graduate GPA). Thus, it is important that test users consider the limitations related to the samples (data) in the predictive validity study.

**Range restriction.** As was discussed above, restriction in range of scores limits the magnitude of the predictive validity. This seems to be particularly true with GRE studies in which the range of GRE scores of the individuals accepted in graduate programs are generally smaller due to a required minimum score, excluding the lower scores. This means that the variance (difference) of the scores is small, and the magnitude of the correlation may be small, as a result.

Furthermore, the range of the criterion measures used in GRE studies is also restricted. For example, one of the most frequently used criterion measures in GRE studies is graduate GPA. Because most graduate programs require students to maintain the minimum GPA of B, and most students' GPAs are within this range, the range of graduate GPAs (criterion scores) is restricted to narrow range from a B to an A (3 to 4

points). Consequently, this restricted range of criterion scores reduces the magnitude of the validity coefficients. This means that for validity studies of GRE scores predicting success in graduate school need very large samples to detect the correlations.

The criterion. The predictive validity of a test score is defined by its relationship to the criterion variable, and the criterion variable has direct bearing on the magnitude of the validity. In the GRE validation study, the criterion is success in the first year of graduate school. This criterion is a concept and cannot be measured directly (Crocker & Algina, 986). For this concept to be measured, an operational measure (criterion measures) is needed. The criterion measures used in the GRE study are several. While a variety of variables can be used as a criterion measure, certain characteristics are desired for a variable to be considered an adequate indicator of the criterion.

The three most desirable characteristics of a criterion measure are the following: First, a criterion measure should be relevant to the concept of the criterion (Brown, 1983; Mehrens & Lehmann, 1987; Willingham, 1974). This requirement seems obvious because if the criterion measure does not measure the concept adequately, it will be of no use. The criterion measure should "reflect the important facets of the conceptual criterion" (Brown, 1983, p. 102).

A second desired characteristic of a criterion measure is that it be reliable (Brown, 1983; Mehrens & Lehmann, 1987; Willingham, 1974). As a measure, criterion scores must be reliable if they are to be useful. But, many GRE researchers addressed problems with criterion measures used in the predictive validity studies. That is, "accurate estimates of the criterion reliability are usually not available" (Breland, 1979, p. 4), and that the criterion measures (e.g., GPA) are unreliable (Kuncel et al., 2001; Willingham,

1974). The unavailability of accurate estimates of the criterion reliability or unreliable criterion measures will result in considerable error (Breland, 1979) in the accuracy of the prediction.

A third desired characteristic is that it be free from contamination (Anastasi, 1976; Brown, 1983; Mehrens & Lehmann, 1987). Criterion contamination occurs in a situation where "the criterion score is influenced by the knowledge of the predictor score" (Mehrens & Lehmann, 1987, p. 81). For example, if a professor assigning the grades knows a student's GRE scores, and the knowledge of the GRE (predictor) scores of the students affects the grades he/she assign for the students, it is considered that the criterion score is contaminated. In other words, the predictions based on the data are likely to be in error.

#### **Operational Definitions of Terms**

**Criterion.** This term refers to the variable that is predicted (Gay, 1996, p. 619) in the present study (Gay, 1996, p. 619). The criterion variable in this study is success in graduate school, as measured by first year graduate GPA, cumulative graduate GPA, and degree attainment status.

**Cumulative graduate GPA (CGPA).** This term is defined differently for master's and doctoral students. Because master's students graduate in shorter period of time, three years on average (NCES, 2007), as compared to doctoral students, six years on average (NCES, 2007), master's students' CGPA is defined as GPAs cumulated over 12 months from the first enrollment in a master's degree program. For doctoral students, it is defined as GPAs cumulated over 24 months from the first enrollment in a doctoral degree program (Kuncel, et al, 2000).

**Degree attainment status.** The variable degree attainment status had two levels, "degree attained" and "degree not attained". Students who attained degrees within the time limit set by the university were included in the category of 'degree attained', and students who either did not attain the degrees within the time limit or dropped out of the program were included in the category of 'degree not attained'. The time limits were 6 years for master's students, and 9 years for doctoral students.

**Differential prediction.** This term refers to a finding where the best prediction equations and/or the standard errors of estimate are significantly different for different groups of examinees "(Young, 2001, p. 4).

**Ethnicity.** This term refers to five categories based on the guidance of definitions for new race and ethnicity categories issued by the U.S. Department of Education. It consists of American Indian or Alaska Native, Asian, Black or African American, Hispanic or Latino, Native Hawaiian or Other Pacific Islander, and White (NCES, 2012). For the purpose of the present study, ethnicity includes seven racial, ethnic, and other groups: African American, Asian American, Hispanic, Native American, White, International students, and Not-Reported groups, in accordance with classification system used by the university. The Not-Reported (NR) group represents students whose race or ethnicity was not reported.

**First year graduate GPA (FGPA)**. This term was defined as the grade point average (GPA) of all courses students took during the first year that was calculated by the university.

**Graduate GPA (GGPA).** This term is used only in references to previous studies, and refers to graduate GPAs cumulated over 2 or more years of course work or the final graduate GPA, as used in previous GRE validation studies (Kuncel et al., 2000).

**Graduate Record Examination (GRE) Test.** This term refers to a standardized test for admission to a graduate school published by the Educational Testing Service (ETS). It consists of GRE General Test and GRE Subject Tests. GRE General Test includes GRE Verbal Reasoning (GREV), GRE Quantitative Reasoning (GREQ), and GRE Analytical writing tests (GREA). GRE subjects (GRES) Tests are tests for eight disciplines: Biochemistry, Biology, Chemistry, Computer Science, Literature in English, Mathematics, Physics, and Psychology (ETS, 2009). For the purpose of this study, however, GRE test includes only GRE verbal, and GRE quantitative reasoning sections of the General Test.

**High-stakes decision.** This term refers to a decision whose result has important consequences for students (U.S. Department of Education: Office for Civil Rights, 2000, p. 75).

**Large university.** This term refers to a university that awards a total of 5, 000-9,999 degrees annually (Florida Department of Education, 2005).

**Predictor**. This term refers to the variable upon which the prediction is based (Gay, 1996, p. 623) in the present study. The predictor variable in this study includes the GRE Verbal Test score, GRE Quantitative Test score, and undergraduate GPA.

**Reliability**. This term refers to the degree to which a test consistently measures whatever it measures (Gay, 1996, p. 624).

**Success in graduate school**. This term was defined as progressing through the programs maintaining the required GPA of B or above, and ultimately attaining the degrees within the time limit (6 years for master's, and 9 years for doctoral programs). Success was measured by three academic variables, first-year graduate grade point average, cumulative graduate grade point average, and degree attainment.

**Test.** This term refers to a standard procedure for obtaining a sample of behavior from a specified domain (Crocker & Algina, 1988, p. 6).

**True score**. This term refers to a score "considered part of the 'true', stable, or unchanging part of a person's observed score" (Mehren & Lehmann, 1987, p. 56).

**Validity.** This term refers to the degree to which accumulated evidence and theory support specific interpretations of test scores entailed by proposed uses of a test (AERA, APA, & NCME, 1999, p. 184).

**Validity coefficient.** This term refers to a correlation between test score and criterion measure (Anastasi, 1976, p. 163), and is expressed in the correlation coefficient,  $r_{.}$ 

### Significance of the Study

The present study is significant in terms of two major perspectives: practical, and research perspectives. From a practical perspective, findings of the present study will provide three benefits to graduate school administrators. First, the findings will provide graduate school administrators with new institution-specific validity data of undergraduate GPA and the GRE scores, which can be used for reference in making admission decisions. Second, the findings will provide graduate school administrators with empirical and professionally defensible evidence to support the current practice of

using undergraduate GPA and GRE scores for admission considerations. Third, findings will provide graduate school administrators with scientific data to review or update educational policies and procedures that involve the use of undergraduate GPA and GRE scores.

From a research perspective, new evidence for differential predictions at graduate study level is of significance as a resource for reference to future GRE validation researchers. Predictive validity of the GRE scores has been of interest since the inception of the use of test scores for admission considerations. Numerous studies have been conducted and most of the studies found consistent but small correlations between the GRE General Test scores (GRE Verbal Reasoning, and GRE Quantitative Reasoning Test scores) and success in graduate school, often measured by first year graduate GPA (FGPA), and cumulative GPA (CGPA).

However, differential predictions have become of interest for test users as the graduate student population has changed, and more minority students have been entering graduate programs. The question has been whether the GRE scores equally predict success in graduate school for these special populations. Existing literature that examined differential prediction is scarce and inconsistent mainly due to a small sample size of minority students. Findings relative to differential prediction of the present study, based on diverse student population (n = 7,367: masters, 5,990; doctoral: 1,377), will be useful as a resource reference for future GRE validation researchers, and add new knowledge to the field.

#### Assumptions

The present study is based on several assumptions: First, "there is a ceteris paribus assumption" (Kane, 2006, p. 56). Ceteris paribus is literally translated as "other things being equal; if all other relevant things, factors, or elements remain unaltered" (Merriam-Webster's Collegiate Dictionary, 2003, p. 203); or, in statistical terms, held constant. That is, to examine the accuracy of a specific prediction made from a test score, it is necessary to assume that all the other relevant factors remain the same, or hold all the other relevant factors constant, so the specific relationship between the test and the criterion is studied exclusively. This is the fundamental assumption in test validation and in use of test scores (Kane, 2006).

Second, graduate students with different background characteristics, for example, different degree level, GREV score, GREQ score, Total GRE scores, and UGPA, may behave differently in their academic endeavor and accomplish varying degrees of success in graduate school as a result.

Third, the predictors and the criterion measures selected for the study were assumed to be relevant to the GRE predictive validity study as these variables are the ones most frequently used in similar studies.

Last, it was assumed that the predictive relationships of the GRE scores to success measures are linear. The linear relationship is assumed because previous validity studies found that the relationships between a test and the subsequent performance measures were "almost invariably linear" (Schmidt & Hunter, 1981, p. 1130) with little evidence for nonlinearity (Coward & Sackett, 1990) both in industrial and in academic settings.

Therefore, the relationships between the GRE scores and the graduate school success measures were assumed to be linear.

#### **Delimitations**

Although it would be ideal to investigate all graduate student records, the present study confined itself to investigating student records in both master's and doctoral degree programs at a large racially and ethnically diverse public university in Southeast Florida. A total of 7,367 graduate student records (masters, 5,990; doctoral: 1,377) from 2000 to 2010 were used to evaluate the relationships between GREV, GREQ, UGPA, and FGPA, CGPA, and degree attainment status.

#### Summary

This chapter included the background to the problem, problem statement, purpose of the research, research questions, conceptual framework, definitions of terms, significance of the study, assumptions, and delimitations. Chapter 2 includes the related literature review. Chapter 3 delineates the research methods employed for this study. Chapter 4 presents the findings of the study. Chapter 5 discusses the findings, implications for practice, and further research, and ends with a conclusion.

#### **CHAPTER II**

#### LITERATURE REVIEW

The purpose of the present study was twofold. One was to determine the degree of relationships among the GRE General Test scores (GRE Verbal Reasoning Test scores [GREV] and GRE Quantitative Reasoning Test scores [GREQ]), Undergraduate GPA (UGPA), and success in graduate school, as measured by first year graduate GPA (FGPA), cumulative graduate GPA (CGPA), and degree attainment status, based on a population that includes a large number of minority students. The other was to determine whether the relationships between a composite scores (GREV, GREQ, and UGPA) and the three success measures (FGPA, CGPA, and degree attainment status) differ by race/ethnicity and sex , more specifically, whether the relationships differ between minority students (African American, Asian American, and Hispanics; note that international students were included as one category of racial/ethnic group in the study) and White, non-Hispanic (White) students, and between men and women.

In this chapter, the conceptual and empirical literatures relative to the validation of the GRE scores were reviewed. "Validation is the process of examining the accuracy of a specific prediction or inference made from a test score." (Cronbach, 1971, p. 443). Thus, the main attention of the present study will be concentrated upon checking the accuracy of the specific prediction proposed by ETS. That is, "the GRE General Test scores are valid predictors of success in the first year of graduate school for all students" (ETS, 2009, p. 7). While the main concern of the present study should be empirically checking on the proposed prediction, the rationale for the proposed prediction needs to be
justified first, i.e., explaining why the GRE General Test scores should predict success in the first year of graduate school for all students.

However, there is no theory that directly explains why the GRE General Test scores should predict success in the first year of graduate school for all students. Thus, it may best be explained by reference to the theories that can be related to the rationale. The concepts and theories that are relevant to the rationale will be discussed in the conceptual review section of this chapter. Following the conceptual review is empirical review. The empirical review focuses on examination of validity evidence (validity coefficients and regression coefficients) from past research studies to check on the proposed prediction.

## **Conceptual Review**

The conceptual review comprises rationales that can be related to the prediction proposed by ETS. That is, "the GRE General Test scores are valid predictors of success in the first year of graduate school for all students" (ETS, 2009, p. 7). But, three items in the prediction need to be clarified before considering the rationales for the proposed prediction. One is "What constitutes General Test scores?" Another is "How is success in the first year of graduate school measured?" and still another is "What exactly does it mean that the GRE scores predict the first year GPA?"

First, with respect to the definition of the General Test scores, by the GRE General Test scores it means the two GRE General Subsection test scores, that is, the GRE Verbal, and the GRE Quantitative test scores (ETS, 2009). Note that there are three subsections in the GRE General Test: GRE Verbal, GRE Quantitative, and GRE Analytical Writing (GREA). But, the GREA is not included. In other words, the two

subsection test scores are the predictors of success in the first year of graduate school for all students.

Second, with respect to the measure for success, the measure for success is the first year GPA in graduate school. Although the ETS has not specified the measure of success expressly, its researchers indicated that the FGPA is the principal success measure (Wilson, 1979; Schneider & Briel, 1990). Thus, the prediction proposed by ETS can be restated in more concrete terms. That is, the GREV and GREQ scores are the valid predictors of the FGPA for all students.

Finally, with respect to the meaning of the GRE scores predicting FGPA, it means that there are positive correlations between the individual GREV and GREQ scores and the FGPA. In other words, those who have higher GREV and GREQ scores tend to exhibit higher FGPA while those who have lower GREV and GREQ scorers tend to exhibit lower FGPA. Thus, what ETS is saying is that because changes in the individual GREV and GREQ scores tend to be associated with changes (increases or decreases) in FGPA, graduate school administrators can predict (estimate using a correlation coefficient and a regression equation) the first year GPA of their applicants by knowing the two subtest scores.

Now that the three items are clarified, the rational for how the GREV and GREQ scores should predict the first year graduate GPA for all students can be considered.

### **Rationales for the Prediction**

The rationale for the prediction proposed by ETS may be derived from answers to the following questions. (a) What is a test? (b) What is a test score? (c) How to interpret a test score? (How can we interpret an observed GREV scores and GREQ scores?) (d)

What are the causes and implications of test scores? (Cronbach, 1971) (e) How are test scores related to the succeeding performance in graduate school (e.g., as measured by the first year graduate GPA)?

The rationales for the two GRE General Subtest scores predicting the first year graduate GPA may be explained by reference to a combination of three theories, a test theory, a performance theory (Kuncel et al, 2001), and knowledge acquisition theory. The test theory may provide answers to the questions, a, b, and c, above that are related to interpretation of test scores, while the performance theory in conjunction with a knowledge acquisition theory may generally answer the questions, d and e, by explaining the causes of the test scores and the implications for succeeding academic performance in graduate school.

### **Test Theory**

GRE is a graduate admission test, and GREV and GREQ are the two of the three subsection tests of the GRE General Test that most graduate schools use in admission considerations. In this section, several questions are considered to provide the definitions of the terms or concepts that are related to interpretation of the test (GRE) scores. Questions included are as follows: What is a test? What is a test score? How to interpret a test score?

### What Is a Test?

A test refers to a measuring instrument used to differentiate among individuals by presenting them "a standard set of questions to be answered" (Mehrens & Lehmann, 1987, p. 7). The result, the data, gathered from a test is often used to make various decisions that may affect individuals' personal and professional lives in a significant way.

A test under consideration is a psychological measurement, more specifically, standardized psychological measurement. By standardized it means that the test is commercially prepared by experts in measurement and subject matter. A standardized test consists of the same fixed set of questions, and is administered with the same set of directions and timing constraints. The standardized test also has to be scored under uniform conditions to be considered as a standardized test (Mehrens & Lehmann, 1987). By psychological measurement (test), it refers to a standard procedure to measure characteristics or properties of people by obtaining a sample of performance from a specified domain (Crocker & Algina, 1986). While standardized psychological measurement is called a few other names including psychometrics (Nunnally, 1978), mental test (Gulliksen, 1967) and test, the word, test, will be used in this section to simplify the language.

A test is broadly classified into two categories depending on the type of performance a test samples (Crocker & Algina, 1986; Mehrens & Lehmann, 1987; Singleton, Straits, & Straits, 1988). They are (a) aptitude and achievement tests (cognitive test), and (b) interest, personality, and attitude inventories (noncognitive test). In aptitude and achievement tests such as the GRE, examinees are instructed to do their best so that the results can reflect individuals' maximum performance. In interest, personality, and attitude inventories (noncognitive measures), respondents are asked to report their typical feelings, attitudes, interests, or reactions to situations so that the results can reflect the typical psychological traits of the individuals.

While the term test is designated only for cognitive measures, i.e., aptitude and achievement tests (Mehrens & Lehmann, 1987), the term inventory is used for

noncognitive measures. Because the main concern of the present study is GRE aptitude test (Pressley & McCormick, 1995; Zwick, 2002), the term test signifies an aptitude and achievement test (cognitive test) in this section unless otherwise specified.

# What Is a Test Score?

A test score (an observed test score) is the simple or weighted sum of the correct answers taken from a set of questions that were presented to the examinees (Gulliksen, 1967). That is, from an examinee's answers the characteristics or properties of the examinees in a specified domain are measured in terms of a numerical value. Each answer is checked against correct answers and to each a numerical value is assigned: One or more points are assigned to each correct answer and zero to each incorrect answer (Gulliksen, 1967). In general, a test yields a range of scores as the responses of the examinees may vary.

Test scores are used to differentiate among people and to make decisions (Mehrens & Lehmann, 1987). But, to use test scores, test users first have to interpret the test scores. While test scores are to be interpreted, it is helpful to know that a given test score (an observed test score) does not necessarily represent the actual ability of a person. That is, an observed test score (X) consists of two components, X = T + E, where X = an observed score, T = true score, E = error score (Gulliksen, 1967). T represents the portion of the observed score that is considered relatively stable or unchanging in a set of tests. On the other hand, error, E, is "the discrepancy between an examinee's observed test users score and his or her true score" (Crocker & Algina, 1986, p. 110). So, the test users would be interested in determining the true score from the given observed score first.

True score can be determined using the equation, X = T + E, if we know the numerical value for error (*X*, an observed score is given). That is, if the two values (an observed score and the error score) are known, the estimate for the true score will be the difference between the observed score and the error score, T = X - E. An error score can be estimated based on a test theory. The assumption is that the mean error averages to zero,  $M_E = 0$  (Gulliksen, 1967, p. 6). This assumption is based on another assumption, i.e., the errors in a test are random. While there are two basic types of errors in measurement, random or chance errors, and constant or systematic errors (Gulliksen, 1967; Mehrens & Lehmann, 1987; Singleton et al., 1988) systematic errors are usually ignored. Instead, only the random errors are considered in test scores. Systematic errors that occur in test scores do not result in variations because the errors are systematic. Therefore, only the random errors that cause variations in test scores are considered in the measurement.

Random errors may occur for many reasons, including the questions in a test, the personal issues of the test takers, and temporary variations in the test administration. "an ambiguously worded question will produce random errors by eliciting responses that vary according to respondents' interpretations of the question's meaning" (Singleton et al, 1988, p. 113); the test takers' personal issues, including motivation, mood swings, careless marking, and guessing "may cause the person sometimes to answer correctly an item that he does not know, and sometimes to answer incorrectly an item that he does know" (Gulliksen, 1967, p. 5); and such things as "any change in directions, timing, or

amount of rapport with the test administrator could cause score variability" (Mehrens & Lehmann, 1987, p. 55).

With respect to  $M_E = 0$  (Gulliksen, 1967, p. 6), the mean error is assumed to be 0 (zero) because if we give a person the same test many times and obtain a large number of scores "the positive and the negative errors [in the observed scores] will cancel each other, and the mean error will be zero (Mehrens & Lehmann, 1987, p. 56). While it may not be correct to say that errors are equal to zero "In actual practice, it is customary to assume that the equation ( $M_E = 0$ ) holds exactly for any particular sample that is being considered" (Gulliksen, 1967, p. 7). This may mean that when we do not know the exact value for an error score we would do best to use 0 (zero) as the estimate for the mean error.

Now that it is assumed that the mean error score is equal to zero, we can rewrite the estimate for the true score as T = X - 0, and we get T = X or X = T. In this equation, Tis the mean of the true score while X is the mean of the observed score. That is, because the error score is the mean of errors  $M_x$  in many observations, the observed score, X, is the mean of the observed scores,  $M_x$ , and the true score is the mean of the true scores,  $M_T$ in the same observations (Gulliksen, 1967). We can see that the mean of the observed test score is the mean of the true score. This means the mean of the observed test score is the true score because "theoretically, the true score of an individual does not vary" (Mehrens & Lehmann, 1987, p. 57).

So far, two assumptions relative to observed test scores have been explained: One is that an observed score consists of two components, i.e., true and error scores; and the other is that the mean of the observed scores is the true score. These assumptions are

useful in interpreting test scores, which will involve estimating the true scores of a given observed score and differentiating people based on test scores.

#### How to Interpret a Test Score.

Test scores have frequently been interpreted using the normal curve model when the test scores are normally distributed (Brown, 1983). "The normal curve is useful because in a normal distribution there are specifiable relationships between standard (z) scores (the number of standard deviations from the mean) and the proportion of scores falling in various areas of the distribution" (Brown, 1983, p. 480). That is, if the scores of a group of examinees are normally distributed, the mean plus 1 standard deviation (+ 1 *SD*) will include about 34% of all the observed scores, and the mean minus 1 standard deviation (-1 *SD*) will include about 34% of all the observed scores. So the mean  $\pm 1$  *SD* will include about 68% of all the observed scores. Or, the mean plus 2 *SD* will include about 47% of all the observed scores, and the mean minus 2 *SD* will include about 47% of all the observed scores (the mean  $\pm 2$  *SD* will include about 47% of all the observed scores (the mean  $\pm 2$  *SD* will include about 95% of all the observed scores).

The z score is a type of transformed score that expresses how far a score is from the mean in terms of the number of standard deviations (Brown, 1983; Gay, 1996). In other words, the z-score is the number of standard deviations from the mean. If a set of scores is transformed into a set of z scores, the mean is z of 0, and a score which is 1 *SD* above the mean is z of 1; a score which is 1 *SD* below the mean is z of -1 (Gay, 1996, p. 443).

The normal curve model can also be used in estimating the true score of a given observed score and differentiating people based on test scores (Brown, 1983). With

respect to estimating the true score of a given observed score, it is assumed that the person's observed scores will be distributed normally around the person's true score. That is, because the mean of an individual's observed scores is the person's true score, test scores (in this case, the person's observed scores) will be distributed normally around the mean (in this case, the person's true score). Thus, using the areas of the normal curve, one can say that the person's observed scores will fall between  $\pm 1$  standard deviation ( $S_e$ ) of the person's true score approximately 68% of the time, or  $\pm 2 S_e$  of the person's true score about 95 % of the time.

However, note that the standard deviation used in calculating the interval is the standard error of measurement ( $S_e$ ), not the standard deviation (SD). The normal curve model can be used with the standard error of measurement ( $S_e$ ) because the standard error of measurement is the standard deviation of the distribution of observed scores around an individual's true score (Brown, 1983, p. 90). In other words, the standard error of measurement ( $S_e$ ) is the standard deviation of a person's observed scores. Therefore, the true score of an individual is estimated using the standard error of measurement.

Suppose, for example, that an observed GRE Verbal score of a person is 500 and  $S_e$  is 34 (ETS, 2009). Then, one can say that the probability is .68 that a given observed score (in this case, the GRE Verbal score of 500) will fall between  $\pm 1 S_e$  of the true score, and .95 that an observed score will fall between  $\pm 2 S_e$  of the true score. In other words, we can be about 68% confident that the person's true score lies between 500  $\pm$  (1) 34, or 466 to 534 and 95% confident that the person's true score lies between 500  $\pm$  (2) 34, or 432 to 568. Now that we can estimate the true scores of a given observed score, the next task is to differentiate among people based on these test scores.

The normal curve model can also be used to differentiate people based on the test scores. That is, z scores and the proportion of scores falling in various areas of the distribution are used to differentiate people. To differentiate people, test scores of individuals are transformed into z scores, and percentile rank equivalents for the z scores are obtained. "The percentile rank for a score is defined as the percentage of persons in the norm group who obtain lower scores" (Brown, 1983, p. 158), and indicates the person's relative ranking (position) in the specified norm group. Norm group refers to a representative sample of a specified population used as the basis for interpretation of individual test scores.

Suppose, for example, that a person's GREV score is 578 or z score of +1 (the mean of GREV was 457, and the *SD* was 121, ETS, 2009). Then, the percentile rank for the score is placed on the 84 percentile with respect to the norm group (the norm group consists of all examinees who took the test between July 1, 2005, and June 30, 2008, ETS, 2009). That is, because the z score for the mean of the GREV of 457 is 0, and the percentile rank equivalents for the mean is placed on the 50<sup>th</sup> percentile, the percentile rank equivalents for the 578 is the 84<sup>th</sup> percentile (one standard deviation above the mean, i.e., 50 percentile + 34 percentile). This indicates that 84 percent of the people in the norm group obtained lower scores, or the person scored better than 84 percent of the norm group. So, percentile ranks differentiate people based on the percentage of people in the norm group who obtain lower scores. And, this is the method used by ETS, the GRE test publisher.

So far, terms and concepts relative to interpreting test scores have been explained. Because one of the major tasks of this chapter is to examine the rational for why the

GREV and GREQ scores should predict the first year graduate GPA, understanding the relevant terms and concepts was a prerequisite to the task.

In the next section, concepts relative to the rationale for why the GRE General Test scores should predict success in the first year of graduate school for all students are summarized. The summary centers about explaining the relationship between the GRE General Test scores and subsequent performance in graduate school. While there is no theory that can directly explain the relationship, a knowledge acquisition theory (Gagne, 1962) integrated into a performance theory (Campbell, McCloy, Oppler, & Sager, 1993) can illuminate the related relationships. The related relationships include two distinguishable relationships and three events. The two relationships are (a) the relationship between the causes of the test scores and the test scores; and (b) the relationship between the test scores or performance). The three events include (a) the causes of the test scores (determinants of performance); (b) test scores (performance); and (c) subsequent performance in graduate school (implications of performance). The related theories are described in the following section.

#### Theory of Performance and Knowledge Acquisition

Test scores (e.g., GRE) represent performance of the test-takers. Performance is idefined as the execution of an action or behavior (Merriam-Webster, 2003, p. 920), but it is further defined as the execution of actions or behaviors that are directed toward the achievement of goals of an individual or organization (Campbell et al., 1993). Some of these actions and behaviors are directly observable while others are not. Actions or behaviors involved in answering questions on a test (cognitive behaviors) are not directly

observable. Nonetheless, all of those actions or behaviors regardless of observability are considered performance as long as they are relevant to the achievement of goals. In this context, individuals taking tests is performance, and their test scores are the representations of the individuals' performance.

Performance is involved in various tasks or jobs. If performance is involved in occupational setting, it is regarded as job performance, and if performance is involved in educational setting, it is regarded as learning performance (for a learner). Thus, there are various types of performance. In addition, performance is multidimensional, such that for each performance there are a number of performance components. "The performance components are the distinguishable categories of things people are expected to do in a job" (Campbell et al., 1993, p. 42). For example, college students whose goals include obtaining a degree are expected to do many different categories of things (performance component) in their endeavor to achieve the goals. Some examples of the performance components include establishing a study schedule with an advisor, writing papers for a course, presenting data in a statistics class, etc. While the students will carry out similar performance components, their performance on each performance component is likely to differ from each other.

Accounting for individual differences is not simple. But "individual differences on each specific performance component are viewed as a function of only three major determinants - declarative knowledge, procedural knowledge and skill, and motivation" (Campbell et al., 1993, p. 43). And, this model applies to both occupational and educational settings (Campbell, 1993; McCloy et al., 1994). In other words, these three types of capabilities account for individual differences in various performances across the

board. This means that individuals' performance on a test can also be accounted for by these three determinants.

However, because the GRE is a measure of cognitive abilities, the motivational determinant is assumed to be equal. Thus, only the declarative knowledge, and the procedural knowledge and skill that are relevant to the present study will be discussed below.

One of the major performance determinants is declarative knowledge. The declarative knowledge also known as verbal information (Gagne & Medsker, 1996, p. 31) is knowledge about facts, names, and labels (Campbell et al., 1993; Gagne & Medsker, 1996). A person who has acquired declarative knowledge is able to state, describe, and recall the facts, names, and labels as was initially presented (sometimes, not identical but similar) to her or him. A person who is equipped with this kind of knowledge understands what is to be done about the given task and performs better than those who are not. Thus, the declarative knowledge is considered a prerequisite for various performances (Campbell et al., 1993; Gagne & Medsker, 1996; McCloy et al., 1994) whether it is for work or learning. Acquisition of this kind of knowledge is largely dependent upon repeated practice in recalling the information (Gagne & Medsker, 1996).

Another major performance determinant is procedural knowledge and skill. The procedural knowledge and skill also known as intellectual skills (Gagne & Medsker, 1996) include four types of hierarchically related cognitive skills. The four types of hierarchically related cognitive skills are discrimination skills, the skills capable of applying concepts, the skills capable of applying rules, and the skills capable of applying higher order rules. These skills are hierarchical and cumulative in their nature. By

hierarchically related, it means that these skills are positioned at four different levels in the skills hierarchy, and the skills in each level are subordinate to the skills above. More specifically, the discriminations skills are at the bottom of the hierarchy, and the skills capable of applying concepts, rules, and higher order rules are in ascending order. That is, above the discriminations skills are the skills capable of applying concepts. Above that (the skills capable of applying concepts) is the skills capable of applying rules, and above that (the skills capable of applying rules) is the skills capable of applying higher order rules.

"The discrimination skills are defined as the ability to distinguish different physical phenomena, such as sounds, shapes, colors, or textures" (Gagne & Medsker, 1996, p. 32). These skills are characterized as the simplest skills of the four intellectual skills. The acquisition of these skills relies on auditory, visual, or kinesthetic senses. The next higher level (or more complex) skills are the skills capable of applying concepts. The skills capable of applying concepts are defined as the abilities "to classify phenomena using critical attributes" (Gagne & Medsker, 1996, p. 32). These skills are characterized as more complex skills than the discriminations skills. The acquisition of these skills relies on individuals' previously acquired discriminations skills. The next higher level (or more complex) skills are the skills, capable of applying rules. The skills capable of applying rules are defined as the abilities "to specify relationships among two or more concepts" (Gagne & Medsker, 1996, p. 32). These skills are characterized as more complex skills than the skills capable of applying concepts. The acquisition of these skills is dependent upon individuals' previously acquired skills that are capable of applying relevant concepts. Finally, the skills, capable of applying higher order rules are

defined as the abilities "to combine multiple rules to perform a task or solve a problem" (Gagne & Medsker, 1996, p. 32). These skills are characterized as the most complex skills of all of the four intellectual skills. These skills are considered the most critical skills as they enable a person to solve a given problem (by using several rules relative to the task). The acquisition of these skills requires the presence of the previously acquired skills that are capable of applying rules.

In sum, the intellectual skills (or the procedural knowledge and skill) are hierarchical and cumulative (Gagne, 1962). That is, acquisition of the intellectual skills occurs hierarchically from the bottom up, i.e., lower level skills to higher level skills, or simple to increasingly complex skills, by incorporating lower level (or simpler) skills into increasingly higher level (or more complex) skills.

The characteristics and the acquisition of the two major performance determinants (declarative knowledge, and procedural knowledge and skill) have been explained by a combination of the two theories, i.e., knowledge acquisition theory and performance theory. The characteristics and the acquisition of the two performance determinants were of concern for the present study as they can help explain the causes and the implications of the GRE scores (test performance).

With respect to the causes of the GRE scores, ETS stated that the two GRE General Subtests measure the following:

The verbal reasoning section tests the ability to analyze and evaluate written material and synthesize information obtained from it, to analyze relationships among component parts of sentences, and to recognize relationships between words and concepts. In each test edition, there is a balance among the passages

across three different subject matter areas: humanities, social sciences, and natural sciences. The quantitative reasoning section tests basic mathematical skills and understanding of elementary mathematical concepts, as well as the ability to reason quantitatively and to solve problems in a quantitative setting. There is a balance among the questions requiring arithmetic, algebra, geometry, and data analysis (ETS, 2009, pp. 3-4).

Judging by the description of the GRE, it is obvious that the test is designed to measure the skills that fall into the types of declarative knowledge (or verbal information) and procedural knowledge and skill (or intellectual skills). Because the two GRE General Subtests are measures of declarative knowledge and procedural knowledge and skill, and because these capabilities are acquired only in hierarchical order, the GRE scores indicate the level of examinees' achievement for these two types of capabilities. So it can be stated that the declarative knowledge and procedural knowledge and skill individuals had (acquired) at the time of the test were in part the causes of the GRE scores. It seems that the rationale for the suggestion ETS makes to use GRE scores in admission considerations is based on the view that individuals' performance is in part a function of declarative knowledge, and procedural knowledge and skill.

The two major performance determinants, declarative knowledge, and the procedural knowledge and skill, have been discussed to explain the relationship between the (GRE) test scores and the causes of the test scores. The next task is to consider the implications of the test performance (GRE scores).

What are the implications of the test performance (GRE scores)? In other words, what are the relationships between the GRE scores and subsequent graduate school

performance (e.g., first year graduate school GPA)? The question can be answered similarly. That is, the subsequent graduate school performance of a person will be dependent upon the person's current level of capabilities relative to the subsequent graduate school performance. Because performance is a function of the three determinants (declarative knowledge, procedural knowledge and skill, and motivation), and because these capabilities are acquired only in hierarchical order, the cognitive skills demonstrated in the GRE test will have influence on the subsequent learning performance (e.g., the first year graduate GPA). Examinees who received higher GRE scores are believed to have acquired more skills relevant to graduate study than those who received lower scores. Assuming that the tested domains in GRE are related to those in the first year of graduate school, those who received higher GRE scores will do better in the subsequent performance in graduate school (e.g., first year GPA).

In addition, if we extend the theories beyond the first year, we can predict that all other things being equal it is highly likely that those who do well in the first year of graduate school will do well in the second year of graduate school, i.e., those who do well in the first year may have acquired (cumulated) more skills that are considered prerequisites (required) to the second year of the graduate study, and so on. In contrast, those who received lower scores theoretically will have difficulty in the subsequent learning performances because they lack the required sets of prerequisite skills. Theoretically, those who received lower scores will be able to catch up if they master the required sets of skills, prerequisites to the subsequent learning performance (Gagne, 1962). In short, the implications of the test performance (GRE scores) include that there

is a correlation between the test performance (score) and the subsequent graduate school performance (e.g., first year graduate GPA).

In this section, three theories, a test theory, a performance theory, and knowledge acquisition theory were discussed in relation to the rationale for why the two GRE General Subtest Scores should predict the subsequent learning performances, e.g., first year GPA at graduate school. In the following section, the GRE validity data from past research studies are reviewed to check the accuracy of the specific prediction proposed by ETS. That is, "GRE scores are valid predictors of success in the first year of graduate school for all students" (ETS, 2009, p. 7).

# **Empirical Review**

This section begins with a discussion of the variables that have been used in the past GRE research studies because they had to be examined first to select the related literature. Followed are the outline for the review and the criteria used for selection of the related literature. Then, the validity evidence from the selected literatures was presented. For ease of reading, the validity evidence from the selected literatures is organized by ethnic and sex groups. This section concludes with a discussion of the present study to address the gap in the existing GRE validity study and a summary.

#### Variables: Predictors and Criterion (Success) Measures

The variables involved in the present study are the GRE General Test scores (GREV and GRE Q), undergraduate GPA (UGPA), and success in graduate school, as measured by first year graduate GPA, cumulative graduate GPA, and degree attainment status. They are called a predictor, not independent variable, and a criterion, not dependent variable, because the principal concern of the study involves investigation of

the relationship, not the influence of one variable on another (Keppel & Zedeck, 1989). A predictor is "the variable upon which the prediction is based" (Gay, 1996, p. 623), and a criterion is "the variable that is predicted" (Gay, 1996, p. 619) in a prediction study. Hence, the GRE General Test scores (GREV and GRE Q), and undergraduate GPA are the predictors, and success in graduate school, as measured by first year graduate GPA, cumulative graduate GPA, and degree attainment status, are the criteria for the present study.

As is indicated in the statement by ETS that "GRE scores are valid predictors of success in the first year of graduate school for all students" (ETS, 2009, p. 7), the predictor variables in GRE studies are the GRE scores, and the criterion variable is success in the first year of graduate school. However, ETS researchers (Wilson, 1979; Schneider & Briel, 1990) specified that the principal criterion measure is first year graduate GPA, and used the variable as the criterion measure for GRE studies. Thus, it seems appropriate that the predictive validity of GREV and GREQ scores be checked against the first year graduate GPA.

Nonetheless, researchers have used other predictors and success measures in GRE validity studies than those specified by ETS. For example, some researchers used the total GREV and GREQ scores as a predictor, instead of the individual GREV or GREQ scores, and others included UGPA in the GRE validity studies. UGPA has been used separately and in combination with the GRE General Test scores as a predictor of success in graduate school. In addition, long-range success measures such as cumulative graduate GPA (CGPA) and degree attainment have also been used as criterion measures of success in graduate school.

While it seems inappropriate (even unfair to ETS, the test publisher) for a researcher to use different predictors and different criterion measures other than the proposed predictors and success measures, the use of other predictors and success measures is justified if one considers the practical use of the test scores by graduate school administrators. For example, many graduate schools currently use the combined GRE scores (GREV and GREQ) for admission considerations. Thus, the current practices in many graduate schools should justify the use of the total score (as a predictor) in the GRE validity studies. With respect to the use of UGPA as a predictor, similar logic can apply. Many researchers have used UGPA in the GRE validity studies, and the study results indicated that UGPA is a predictor for success in graduate school. For this reason, use of UGPA as a predictor in the GRE validity study, whether separately or in combination with the GRE scores, is warranted.

On the other hand, researchers have frequently used (cumulative) graduate GPA (CGPA) and degree attainment as success measures for the GRE validity studies. While CGPA is the most commonly used criterion measure, degree attainment is also used in the research studies. The rationale for using CGPA and the degree attainment as a criterion measure can be explained by a reference to the ultimate measure of success in graduate school. That is, when an admission is offered to an applicant, graduate school administrators expect that the applicants would make steady progress and finishes the program successfully. In this case, CGPA can be used as an interim measure of success, and the status of degree attainment can be used as an ultimate measure of success in graduate school (Willingham, 1974).

Thus, the use of long-range criterion measures such as CGPA and degree attainment in the GRE validity study should be reasonable. In addition, the investigation of the relationship between the GRE scores and the long-range criterion measures can be worthwhile as "the accumulation of long-term follow-up studies will indicate the extent to which various types of early criteria are related to long-run outcomes of importance to educators" (Cronbach, 1971, p. 488).

In sum, the predictors and criterion (success) measures in the GRE validity study may include variables that have been used by the researchers. That is, combined GREV and GREQ, UGPA, or a composite of GREV, GREQ, and UGPA may be used for a predictor, and cumulative GPA and degree attainment as a success measure, in addition to the proposed variables (individual GREV and GREQ for a predictor and first year graduate GPA for a criterion measure).

# **Outline for the Empirical Review**

The present empirical literature review focuses on gathering validity evidence (validity coefficients and regression coefficients) from the previous GRE validity studies and examining the proposed prediction, i.e., whether the two GRE General Subtest scores in fact predicted the (first year) graduate GPA for all students. More specifically, once the validity data (validity coefficients and regression coefficients) were collected from the selected empirical literature, the validity data from the literature were examined to check the proposed prediction by ETS in terms of three items. The three items included the existence of the relationship between the GRE scores and the criterion, the degree of the relationship between the GRE scores and the criterion, and similarities or differences in the magnitude of the relationships across different groups of graduate students.

The three items were examined in the following manner. First, the existence of the relationship between the GRE scores and success in graduate school was examined in terms of validity coefficients (r), regression coefficients (B) or standard regression coefficients ( $\beta$ ) reported in the research study. The existence of the relationship was confirmed if the r, B or  $\beta$  were positive and statistically significant. In this case, it was stated that the study found that the GRE scores predicted the (first year) graduate GPA.

Second, the degree of the relationship between the GRE scores and success in graduate school was judged by the magnitude of validity coefficients (r or R), regression coefficients (B) or standard regression coefficients ( $\beta$ ) reported in the research studies. The degree of the relationships was judged principally using r,  $r^2$ , R, and  $R^2$  as most studies use these indices in reporting their validity data. The data was analyzed using the convention. i.e., r or R of .10, .30, and .50, regardless of sign, were interpreted as small, medium, and large coefficients, respectively (Cohen, 1988). Although "r is the more appropriate measure of association, as compared with  $r^2$  in most psychological investigations (Ozer, 1985), the square of the correlation,  $r^2$  or  $R^2$ , were also used as some researchers used  $r^2$  or  $R^2$  as the coefficient of determination. In this case, it was interpreted as the amount (%) of variance accounted for by the predictor(s).

Last, similarities or differences in the magnitude of the relationships were determined by examining the *r*,  $r^2$ , *R*,  $R^2$ , adjusted  $R^2$ , *B*, and  $\beta$  between different groups. If the validity data are similar, it was stated that the study found that the GRE scores predicted the (first year) graduate GPA similarly for the different groups, or vice versa.

### **Criteria Used in Selecting the Related Literature**

Numerous GRE validity studies have been conducted by various institutions (graduate schools and departments), individual researchers, and the test publisher, ETS, since the first use of the test, 1937 (Zwick, 2002, p. 19). However, only the literatures relative to the purpose of the present study and the research questions were selected. That is, only those research studies that investigated the predictive relationships of individual GREV and GREQ to the three criterion measures (first year graduate GPA, cumulative graduate GPA, or degree attainment), and that reported the validity coefficients (and regression coefficients) are considered for the review.

The related literature was searched online through the aid of PsycInfo, PsycArticles, Applied Social Sciences Index and Abstracts (ASSIA), and the ETS website. In addition, relevant articles referenced in the research studies were searched manually. Twenty eight published GRE validity studies (seven meta-analyses, one large scale study, and 20 locally conducted institutional studies) were found for the period of 1965 to 2010. Of the 28 studies, 12 studies that are relevant to the present study were selected, i.e., seven meta-analyses based on White student majority (four studies carried out by the independent researchers; three by the ETS researchers); one large-scale study by an ETS researcher (White student majority), and four institutional studies (three studies on minority student samples, and one on a sex group sample). In short, all metaanalyses and one large scale study were selected to examine the validity data on White student majority while the four studies were selected to examine the validity data on African American, American Indian/Alaska Native, Hispanic, and sex groups.

The reasons are threefold. First, the reason to select the meta-analyses on the White student majority is that the results of meta-analyses on GRE can provide one with adequate information so the person can develop a general conclusion about the predictive validity of the GRE test scores for typical graduate student population as the meta-analysis is a statistical approach that summarizes the results of many studies of basically the same problem (Gay, 1996).

Second, an interest of the present study includes examining the predictive relationships of the GRE scores to success in graduate school across different groups. Thus, it is important that the related literature review includes those studies that involved various different populations and subpopulations.

Third, one large-scale validity study is included as it is considered an important study in the development of the GRE validity study. It is important that the review of the related literature includes the relevant historical developments to the study (Merriam & Simpson, 1995).

## **Organization of the Selected Literature**

The selected literature is organized by ethnic and sex groups with chronological order. However, literatures on White students (seven meta-analyses and one large scale research study) were reviewed first as those comprise most of the research studies. Followed are the research studies based on minority and sex groups. The literatures on minority groups are organized in alphabetical order, i.e., African American, American Indian/Alaska Native students, and Hispanics. The literature on men and women were reviewed last as only a minimum of data was available. For the sake of classification, those studies that did not specify the ethnic composition of the population are assumed to

be composed of a majority of White students (White students are the majority of the graduate student population, CGS, 2008).

#### Validity Evidence Based on White Student Majority

Thacker and Williams (1974) summarized 12 GRE validity studies that were published during the period of 1957-1970. The studies were based on master's and doctoral student data from Education and Psychology department samples (N = 1,874). Predictors included individual, and combined GREV and GREQ scores. Criterion measures included both short- and long-range criteria, i.e., graduate GPA on various time dimensions in master's and doctorate degree levels, doctoral comprehensives, faculty rating, graduated/not-graduated, and success after six years.

The results indicated that although the validity coefficients varied greatly, the GRE scores were valid predictors of the (first year) graduate GPA in most of the studies. That is, the relationship between the GRE scores and the (first year) graduate were positive and significant (the existence of the true relationship). The correlation coefficients (*r*) of the GREV or GREQ with the first year GPA or graduate GPA (including the first semester graduate GPA) were small to medium, ranging from .18 to .49 for GREV, and .15 to .37, for GREQ.

In discussing the findings, the researchers initially considered the range restriction in graduate GPA as a factor that reduced the correlations between the GRE scores and graduate GPA. But they found that the correlations for other criterion measures were also small. Consequently, the researchers concluded that "the weight of the evidence in these articles suggests that the wide usage of the GRE as a selection instrument must be questioned" (p. 943).

It is true that restriction in range of scores reduces the validity (correlation) coefficient (Hartnett & Willingham, 1980) as "the magnitude of a correlation depends on the range of individual differences - that is, the variance of the predictor and criterion scores" (Brown, 1983, p. 104). This means that the magnitude of a correlation is small when the variance (difference) of the predictor and criterion scores is small, or vice versa. More specifically, if the range of individuals' scores is restricted (individuals' scores are not so different), the variance of the scores will be small, and this will influence (reduce) the magnitude of the correlation. Thus, the use of graduate GPA as a criterion measure may have reduced the validity coefficients. Because most graduate programs require students to maintain the minimum GPA of B, and most students' GPAs are within this range, the range of graduate GPAs (criterion scores) is restricted to Bs and As (or 3 and 4 points), excluding Cs, Ds, and Fs (or 2, 1, and 0). Hence, when the graduate GPA is used as the criterion score, the range of scores is restricted. This may have reduced the related validity coefficients.

With respect to the small to medium correlations between GRE scores and other criterion measures, similar factor (range restriction) could have influenced the results. If the individuals' scores in the sample, both the predictor and criterion score, are not so different, the variances are small among the individual scores, and that can reduce the magnitude of the correlation.

Willingham (1974), an ETS researcher, reviewed 43 GRE validity studies that were published from 1952 to 1972. The study involved 138 independent sets of data from nine fields and included 21,214 graduate student data (N = 21,214). The nine fields in the study were biological science, chemistry, education, engineering and applied science,

English, mathematics, physics, psychology, and social science. The study examined several relationships between predictors and criterion measures. Predictors included GREV, GREQ, GRES (GRE Subject Tests, formerly called the Advanced Tests), GRE composite (V+Q+S), UGPA, and the composite GRE and GPA (V+Q+S+U) scores. Criterion measures included graduate GPA, overall faculty rating, departmental examination, attainment of Ph.D., and time to Ph. D.

The results indicated that the GRE scores were valid predictors of the graduate GPA. The validity coefficients (r) of the GREV, GREQ, GRES, GRE composite (V+Q+S), UGPA, and the composite GRE and GPA (V+Q+S+U) scores with the graduate GPA were .24, .23, .30, .33, .31, and .45, respectively. The graduate GPA had the highest correlation (.45) with the composite GRE and UGPA (V+Q+S+U), and the lowest with GREV (.24). On the other hand, the status of degree attainment has the highest correlation (.40) with the composite GRE and UGPA (V+Q+S+U) and its lowest (.14) with UGPA.

In discussing the findings, the researcher stated that although the use of the composite predictors (GRE and UGPA) improved the prediction, the "accuracy of predicting which students will succeed in a particular graduate school is often no better than modest, especially if such predictions are based only upon a test or a grade record" (p.278). Unreliability of the criterion measures and restriction of range of scores were considered as factors that limit the magnitude of the validity coefficients.

In general, the magnitude of the relationships between the two GRE scores (V+Q) and the graduate GPA similar to those reported by Thacker and Williams (1974). They were in the range of .21 to .32. But, the findings of the study included new information.

That is, the use of the composite score of the GRE and UGPA improved the prediction. This finding suggests that by considering both the GRE scores and the UGPA, one can predict the applicant's subsequent academic performances better. This information is useful for graduate school administrators whose main concern includes improving the educational policies and procedures. This type of research finding has justified the admission requirements of GRE scores and UGPA to graduate schools.

Wilson (1979), an ETS researcher, analyzed student data (N = 4,433) to assess the relationship of the GRE General Test Scores and UGPA to the first year graduate GPA by departments/fields and Verbal and Quantitative fields. Predictors included GREV, GREQ, GRES (GRE Subject Tests, formerly called the Advanced Tests), and UGPA. The criterion measure was the first year graduate GPA (FGPA).

The study involved 19 departments in 39 member institutions of the Council of Graduate schools. The 19 departments/fields comprised biosciences, chemistry, engineering, mathematics, physics, geology, geophysics, economics anthropology, education, English, history, political sciences, psychology, sociology, library sciences, fine arts, music, philosophy, and languages. The student data were pooled from corresponding departments/fields across institutions by department/field and by Verbal (verbally oriented departments) and Quantitative (quantitatively oriented departments) fields.

The results included two types of reports. One is department/field specific validity coefficients and regression coefficients; and the other is the evidence of the association of the GREV scores with the Verbal fields, and the GREQ scores with the Quantitative fields. With respect to the departmental validity and regression coefficients, the results

were reported based on samples that consisted of cases ranging from 82 to 530. The results indicated large variations across the departments. The validity coefficients (*r*) of the individual GREV, GREQ, GRES (GRE Subject Tests, formerly called the Advanced Tests), and UGPA with the first year GPA by department ranged from .05 to .43, .04 to .52, .11 to .54 and .06 to .56, respectively. And, the regression coefficients of the GREV, GREQ, and UGPA with FGPA by department ranged from .08 to .37, .08 to .37, .and .18 to .33, respectively.

Overall, the validity data varied greatly across various different departments/fields and included negative regression coefficients. Variations in department/field specific validity data may occur due to the distinctive characteristics of the various departments/fields or sampling error. While sampling errors can occur due to small department samples, it is the student characteristics that directly affect the magnitude of the validity data. In fact, the student characteristics (as represented in the means and the standard deviations of the students' GRE scores, their UGPAs, and their graduate GPAs) vary greatly across various departments.

With respect to negative coefficients, when they are observed, it indicates the need for more comprehensive investigation to illuminate the particular circumstances involved (Cronbach, 1971; Wilson, 1979). While "there are no priori reasons for assuming that all the coefficients are nonnegative" (Longford, 1991, p. 4), negative coefficients are "perceived as theoretically anomalous" (Wilson, 1979, p. 123).

The findings of the study indicated that only the GREQ was a valid predictor of FGPA (the regression coefficients of FGPA on GREV ranged -.08 to .37, including negative values). In discussing the results, the researcher stated that summarizing

department/field specific validity data is useful to graduate schools because it provides them with more focused validity data for the individual departments/fields, in addition to the overall validity data for the institution. The departmental validity data can help the graduate schools and departments make more informed decisions as they will have more (department-specific) GRE validity information.

With respect to the relationship of the GREV and GREQ scores with Verbal and Quantitative field, the results indicated a "tendency for GREQ to have higher validity than GREV in the quantitative fields (departments), and for GREV to have higher validity than GREQ in the verbal fields (departments)" (p. 20). The median correlations of the GREV and GREQ with the first year GPA in the Verbal field were .31 and .25, and those in the Quantitative field, .20 and .31, respectively. While the finding needs to be cross-validated by other studies it may be interpreted that the verbal and quantitative reasoning skills as measured by the GRE are associated with those required in the related fields of studies.

On the other hand, it implies that students who received high scores in the GREV tend to exhibit higher first-year GPA in Verbal field (verbally oriented departments) while those who received high scores in the GREQ tend to exhibit higher first year GPA in Quantitative field (quantitatively oriented departments). This information is valuable because it provides graduate administrators with empirical evidence for assigning separate weights for the GREV and GREQ scores depending on the applicant's field of study. That is, if the applicant plans to study in a Verbal field, more weight can be assigned for GREV than GREQ in admission considerations, or vice versa.

The study is not a meta-analysis, but it is included because it contributed to the development of the GRE validity study in two significant ways. First, it was the first largest, and most extensive validity study that examined the predictive relationship of the GRE General Test, GRE Advanced Test scores (current GRE Subject Test), and UGPA with the first-year graduate GPA, the principal criterion measure that the GRE test should predict.

Second, a GRE validity study model was developed as a result of the study. That is, the framework or the standards for the current GRE Validity Study Service (ETS) was established by this study (Schneider & Briel, 1990). The standard data needed for a predictive validity study was specified, and a format for the validity result report was established. e.g., using department/field samples as the basic unit of analysis and summarizing the results (validity data) by the departments/fields. The model has been used or adapted by several researchers (Schneider & Briel, 1990; Kuncel, Hezlett, & Ones, 2001, Burton & Wang, 2005) since the study.

In addition, the researcher discussed a need for separate validity studies for various subgroups, including "women, minorities, older students, or foreign students" (p. 4). While more validity data on these subpopulations will be useful to the test users, there is still a need for more research, particularly on minority students as we still do not have as much confidence in the accuracy of the prediction for minority student population as we do for White students (ETS, 2009).

Schneider and Briel (1990), the ETS researchers, analyzed 606 validity studies (*N* = 9,200) to assess the relationship of the GRE Scores and UGPA to success in graduate school. Predictors included the GRE General Test scores (GREV, GREQ, and GREA),

GRE Subject Test scores, and UGPA, and the composite of the predictors (V+Q+A, and V+Q+A+U). The criterion measures included the first year GPA and faculty rating (Note that "beginning with the 1987-88 cycle, a faculty rating scale was predefined as an optional criterion variable" ["Overview," para. 6] of the GRE Test scores; however, no operational definition for the variable was provided in the study; also note that only the variables that are related to the present study are discussed in this review (GREV, GREQ, and UGPA as predictors, and first year graduate GPA, the criterion).

The data were analyzed for all departments combined and five major departments/fields. The five major departments/fields were Natural Sciences, Engineering, Social Sciences, Humanities & Arts, and Education. The minimum sample size for each department was 100 cases (student data).

The results indicated that the GRE scores had moderate correlation with first-year graduate GPA and were the valid predictors of the first year graduate GPA. The overall correlations of the GREV, GREQ, GREA, UGPA, V+Q+A, and V+Q+A+U with the first year GPA were .29, .28, .26, 34, .33, and .43, respectively. While the first year GPA had the highest correlation with the composite score of the GRE and UGPA (V+Q+A+U), the UGPA was the one that contributed most to the highest correlations of the composite predictor (.34). The UGPA, by itself, had higher correlations with the first year GPA (.34) than the GRE General Subtest scores (V+Q+A) combined (.33).

The department/field specific correlations of the GREV, GREQ, GREA, UGPA, V+Q+A, V+Q+A+U, GRE Subject Test (GRE-S), and V+Q+A+U+S with the first year GPA ranged from .22 to .31, .18 to .32, .20 to .28, .29 to .39, .25 to .36, .40 to .47, .21 to .48, and .41 to .60, respectively (Note that the departmental correlations include

additional predictors, the GRE Subject Test scores, and the composite score). In the department/field specific data, the results showed that the first year graduate GPA had the highest correlation with the V+Q+A+U+S (.41 to .60), and the second highest correlation with the V+Q+A+U (.40 to .47). On the other hand, the UGPA was the highest single predictor (.29 to .39).

In discussing the findings, the researchers considered several factors that limited or reduced validity coefficients. Included were unreliability of the predictors and criterion measures, the restriction of range in the predictors and criterion measures, and compensatory selection of students. While unreliability and the restriction of range had been discussed by previous researchers, Schneider and Briel specified compensatory selection of students as an additional factor that decreases observed correlations. Compensatory selection (Dawes, 1975; Schneider & Briel, 1990; Whitworth & Barrientos, 1990) is an admission decision practice where high scores on one admission marker (e.g., UGPA) are allowed to compensate for low scores on another (e.g., GRE score) that is, some students who received low scores on the GRE tests are offered an admission because of their high UGPAs.

While the effects of compensatory selection can be shown as low (or negative) intercorrelations between the GRE scores and UGPA (Dawes, 1975; Schneider & Briel, 1990), the compensatory selection will influence the magnitude of the validity coefficients. That is, the multiple correlations between the multiple predictors (GRE scores and UGPA) and the criterion measure are inflated even when the predictors considered individually have small validity (Dawes, 1975). This means that to properly understand the effects of each predictor on the criterion, it is appropriate to examine the

validity of the predictors separately. This is probably why ETS recommends that researchers use individual scores (GREV, GREQ) for the GRE validity study.

Goldberg and Alliger (1992) reviewed 27 GRE validity studies that were published during the period of 1950 – 1990. The study was based on psychology/counseling department samples (N = 2,754), and the number of subjects in each sample ranged from 23 to 582 with a median of 75.5. Predictors included GREV, GREQ, and GRES (GRE Subject Tests, formerly called the Advanced Tests) scores. Criterion measures included graduate GPA (GGPA), specific course grades, and comprehensive exam. However, only the relationships that are relevant to the present study are discussed in this review, i.e., the relationships among the GREV, GREQ, and GGPA.

The results indicated that the GRE scores were valid predictors of the GGPA, although the magnitudes of the relationships were small. The validity coefficients of the GREV, GREQ, and GRES with GGPA were .15, .15, and .29, respectively while the validity coefficients of GREV, GREQ, and GRES with specific grades were .03, .02, and .01. The validity coefficients of the GREV and GREQ with comprehensive exam were .37, .28 (The validity coefficient of GRES with comprehensive exam was not reported in the study and not included here). The GGPA had the highest correlation (.29) with the GRES while it had lower correlations with GREV (.15), and GREQ (.15).

Overall, the researchers concluded that GRE was a predictor of GGPA, and had fairly low validity. Indeed, the validity coefficients of the study were the smallest that had been reported in meta-analyses that involved GGPA and GREV and GREQ. The correlations were small to medium at best. While the range restriction in graduate GPA

was discussed, the researchers addressed the importance of focus in ultimate measures of success in graduate school such as "Graduation and scientific productivity" (p. 1026).

Morrison and Morrison (1995) reviewed 22 GRE validity studies (N = 5,186) that were published from 1982 to 1993. This study included only those studies that involved the relationship of GREV and GREQ scores with Graduate GPA (GGPA). The researchers considered faculty evaluations and level of postgraduate productivity as a criterion measure but decided to exclude those variables. The reason given for the exclusion of those variables was that the findings of the previous research were inconclusive.

The results indicated that the GRE scores were valid predictors of the GGPA. The validity coefficients of the GREV and GREQ with GGPA were .28 and .22, respectively. Overall, the magnitude of the observed validity coefficients were similar to the previous meta-analyses (except those of Goldberg and Alliger, 1992), most of which were in the range of .20 to .30 for both the GREV and GREQ.

Nevertheless, the researchers stated that "the results of this meta-analysis suggest that the quantitative and verbal components of the GRE possess minimal predictive validity" (p. 311) and concluded that "the average amount of variance in graduate GPA accounted for by performance on these dimensions of the GRE was of such little magnitude that it appears they are virtually useless from a prediction standpoint" (p. 311).

Kuncel, Hezlett, and Ones (2001) reviewed GRE validity studies extensively that had been conducted from the late 1940s to the late 1990s. The study involved 1,753 independent samples and included 82,659 graduate student data (N = 82,659). The predictors included individual and combined GREV, GREQ, GREA, GRE Subject Test

scores (GRES), and UGPA. The criterion measures included first year graduate GPA (FGPA), graduate GPA (GGPA), comprehensive exam scores, faculty ratings, degree attainment, time to complete, research productivity, and publication citation count. The predictive relationships between the five predictors and the eight criterion measures were meta-analyzed in multiple disciplines. However, only those validity data that are relevant to the present study are discussed in this review, i.e., the validity coefficients involving the GRE General Test scores, UGPA, and the three criterion measures (FGPA, GGPA, and degree attainment status).

The data were analyzed for the total group (overall) and four broad disciplines (humanities, social science, life science, and mathematics-physical science). From the outset, the researchers addressed the limitations involved in the GRE validity study, including the range restriction, criterion unreliability and compensatory selection, and corrected the validity coefficients for the range restriction and criterion unreliability. Thus, the result report included both the observed validity coefficients and corrected validity coefficients.

The overall observed (corrected) correlations of the GREV, GREQ, GREA, GRE GRES, and UGPA with FGPA were .24 (.34), .24 (.38), .24 (.36), 34 (.45), .30 (.33), respectively. The FGPA had the highest observed correlation with the GRES (.34), and the second highest, with UGPA (.03). The discipline specific correlations (corrected) of the GREV, GREQ, GREA, GRES, and UGPA with FGPA ranged from .16 to .28 (.06 to .15), .23 to .25 (.02 to .13), .22 to .26 (.04 to .12), 25 to .36 (.00 to .06), and .30 to .31 (.09 to .13), respectively.
The overall observed (corrected) correlations of the GREV, GREQ, GREA, GRES, and UGPA with GGPA were .23 (.34), .21 (.32), .24 (.36), .31(.41), and .28 (.30). The GGPA had the highest observed correlation with the GRES (.31), and the second highest, with UGPA (.028). The discipline specific correlations of the GREV, GREQ, GREA, UGPA, and GRES with GGPA ranged from .21 to .27 (.30 to .39), 18 to .25 (.27 to .38), .24 to .33(.36 to .48), .30 to .37(.40 to .49), and .13 to .38 (.14 to .41), respectively.

The results showed the magnitudes of the overall observed validity coefficients of GGPA with the individual GREV and GREQ scores were similar to those of FGPA with the GRE scores. They were all in the range of .20 to .30. On the other hand, the overall observed validity coefficients of UGPA and the GRES with FGPA and GGPA were consistently higher (medium correlations, about .30) than those of the GREV or GREQ.

The overall observed (corrected) correlations of the GREV, GREQ, GREA, GRES, and UGPA with degree attainment were .14 (.18), .14 (.20), .08 (.11), and .32 (.39), and 12 (.12). The degree attainment had the highest observed correlation with the GRES (.32), and the lowest, with GREA (.08). The discipline specific correlations of the individual GREV, GREQ, GREA, and UGPA with degree attainment ranged from .03 to .41 (.03 to .72), -.07 to .22 (-.09 to .31), -.07 to .37 (-.10 to .49), and. -.02 to .22, (-02 to .22), respectively (please note that the discipline specific correlations of GRES with degree attainment were not reported in the study; and hence they are not discussed here).

In general, both the overall and the discipline specific correlations involving degree attainment with the GRE scores and UGPA were smaller than those observed for the other criterion measures such as FGPA and GGPA. Researchers discussed that "the

differential base rates of graduation from programs may affect the size of the relationship between GRE test scores and degree attainment" (p. 168). While this may be true, there are many other factors that directly influence the person's ability to complete the program and attain the degree, including the personal, professional, and financial reasons (Smallwood, 2004). Given that most graduate students may be involved in complex personal and professional situations during the course of their education, such long range success measure as degree attainment may not correlate well with the academic admission markers, e.g., GRE scores and UGPA used in admission considerations (Dawes, 1975).

While the researchers noted that the results based on subdiscipline be "interpreted with caution as smaller sample sizes compared to the overall analyses result in greater sampling error and less stable estimates" (p. 170), they concluded that GRE and UGPA were generally valid predictors of FGPA, GGPA, and others (comprehensive examination scores, publication citation counts, and faculty ratings). Indeed, the results indicated that the GRE scores were valid predictors of FGPA and GGPA.

Burton and Wang (2005), ETS researchers, conducted the first GRE validity study using computer-based GRE scores that were collected during the 1995-96, 1996-97, or 1997-98 school years (computer-based GRE was introduced in the 1993-94 school year, p. 2). The study involved 1,094 master's and doctoral student data (N = 1,094) from 21 departments in biology, chemistry, education, English, and psychology in seven (7) member institutions of the Council of Graduate schools. The departmental sample sizes ranged from as small as 2 to as large as 453 with a mean of 66. The results of the study were analyzed for all departments combined and for individual departments.

The overall observed (corrected) correlations of the GREV+GREQ, UGPA, and V+Q+U with graduate GPA (GGPA) were.33 (.40), .24 (.32), and .40 (.49), respectively (Note: regression coefficients for the total group were not reported). The magnitude of the overall observed *R* that involved V+Q and GGPA (.33) is similar to that (.31) of Thacker & Williams (1974).

On the other hand, the department specific validity coefficients varied across the departments. The observed (corrected) correlations of the GREV+GREQ, UGPA, and V+Q+U with GGPA ranged from .29 to .39 (.32 to .51), .11 to .29 (.16 to .45), and .38 to .46 (.44 to .62), respectively. The regression coefficients of GREV, GREQ, and UGPA, with GGPA ranged from .056 to .198, .008 to .193, and .021 to .245, respectively. It is noteworthy that all of the departmental coefficients were positive, although some were close to zero. The researchers stated that "no statistical tests" (p. 3) were done in the data analyses.

The results included the validity coefficients relative to the two criterion measures, i.e., GGPA and faculty rating. However, those related to the faculty rating are not included in this review for two reasons. First, faculty rating (on three long-term success measures) is not an established criterion measure, and the quality of the measure has not yet been demonstrated (the measurement process for the variable is still under construction). Second, at this time, the long-term success measures are not the principal criterion measures that the GRE was designed to predict (Schneider & Briel, 1990; Wilson, 1979), thus it not the principal concern of the present study. In addition, the researchers discussed several problems involved in the measurement process of the faculty rating, including unavailability of the information, i.e., "Several departments did

no ratings at all, and most others did not rate all of their students" (p. 38). It seems that these problems may be true with many other graduate schools and programs.

### Summary of the Validity Data Based on White Student Majority

In general, all meta-analyses and the large-scale research study (except departmental data) reported that the GREV, GREQ, and UGPA were valid predictors of FGPA and GGPA for a graduate student population that was consisted of a majority of White students. Most of the studies found that UGPA was a slightly better predictor than the individual GRE scores. The observed validity coefficients for the FGPA were generally similar to those of the GGPA.

In discussions, several factors were considered to account for the low correlations of the GRE scores with FGPA and CGPA. Among them were restriction of range (in both the predictor and the criterion measures), criterion unreliability, and compensatory selection. To rectify the problems, some researchers (Burton & Wang, Kuncel, Hezlett, & Ones, 2001, Schneider & Briel, 1990) statistically corrected the validity coefficients for the restriction of range and the criterion unreliability. While the corrected correlations varied among the predictors and across the different studies (different researchers used different statistical methods), they were generally found to be higher than those of the observed correlations. In addition to these limiting factors, sampling errors might have also occurred in those reports where small department or field samples were used.

In previous studies, the overall observed correlations between GREV and FGPA were small (r < .30) ranging from .22 to .29 while the corrected correlation from one study was medium, .34 (.30 < r < .50). Similarly, the previous studies' overall observed correlations between GREQ and FGPA were also small ranging from .15 to .28, while the

corrected correlation from one study was medium, .38. On the other hand, the previous studies' overall observed correlations of UGPA with FGPA were medium (.30 < r < .50) ranging from .30 to .34, while the corrected correlations from one study was .33.

The overall observed correlations of the individual GREV, GREQ, and UGPA with the GGPA found in previous studies were similar to those of the FGPA. The overall observed correlations between GREV and the GGPA were small to medium, ranging from .15 to .49 (Note that .49 was the correlation based on 24 subjects), while the corrected correlation was medium .34. The overall observed correlations between GREQ and the GGPA were also small to medium ranging from .15 to .37 (The correlation is also based on the same sample of the 24 subjects), while the corrected correlation was medium, .32. The overall observed correlations between UGPA and the GGPA were small to medium ranging from .24 to .31 while the corrected correlations were medium ranging from .30 to .32.

The overall observed correlations between GREV and the degree attainment status from previous studies were small to medium, .14 to .34, while the corrected correlations were similarly small to medium, .18 to .47. The overall observed correlations between GREQ and the degree attainment status were small ranging from .08 to .26, while the corrected correlation from one study was also small .20. Similarly, the overall observed correlations between UGPA and the degree attainment status were small ranging from .12 to .14, while the corrected correlation from one study was .12.

Note that the validity coefficient for the composite of GREV, GREQ, and UGPA (V+Q+U) with the criterion is not summarized here because no comparable data are available.

### Validity Evidence Based on Minority Students

The existence of small to medium relationship between the GRE General Test scores and first-year GPA or graduate GPA for typical graduate students (White) has been known for decades among test users (researchers and graduate school administrators). But what has not been known is whether the known relationship is also true to various different populations and subpopulations, e.g., minority student population (e.g., minority ethnic groups and women). The following subsection includes the published validity data from past research studies whose samples were based on the minority students. It includes a total of four studies.

African American students. Sampson and Boyer (2001) studied the predictive relationship between the GRE scores and the first year GPA (FGPA) at a 'Research I' institution. Data were collected from 160 minority students (N = 160: 144 African American; 13 Hispanic; 2 Native American; and 1 Pacific Islander) who earned a doctoral, specialist, or master's degree during the period of 1988-1997 and who were recipients of a fellowship from the school. On average, the graduate students had UGPA of 3.11; a first year average of 3.51; GREV score of 447, GREQ score of 450; and GREA score of 470. Of these students, 103 were women (64%) and 57 were men (36%). The 96 persons (60%) were conferred master's degrees predominantly in verbal fields, including Social Sciences, 67 (42%), and Humanities, 29 (18%). Predictors included gender, degree conferred (degree), GREV, GREQ, GREA, age, majors, UGPA, and undergraduate institution. The criterion measure was FGPA.

Data were analyzed using regression and correlation methods. Correlations between each of the predictors and the FGPA were computed first, and then the variables

that were significantly associated with the FGPA were utilized in the regression analyses. The correlation analyses indicated that only the GREV score, age, majors, UGPA and undergraduate institution were significantly associated with the FGPA. The correlation coefficients of GREV, age, majors, UGPA and undergraduate institution with FGPA were .39, .21, -.15, .15, and .19, respectively. The correlations of GREQ, GREA, gender, and degree with the FGPA were not significant.

Then, the FGPA was regressed on the significant predictor variables, GREV, age, major, UGPA and undergraduate institution, using a multiple regression method. The adjusted  $R^2 = .25$ , F(5, 144) = 10.05, p < .0001, and approximately 25% of the variability in the first year GPA was predicted from GREV, age, major, UGPA, and undergraduate institution. The majors included five fields: social sciences, behavioral sciences, humanities, biological sciences; physical sciences. The undergraduate institutions were classified into two categories: historically black colleges and universities (HBCU) and predominantly White institution (PWI).

The standardized regression coefficients ( $\beta$ ) for GREV, age, major, UGPA and undergraduate institution were .35, .23, -.16, .18, and .12, respectively. The t - tests indicated that the beta weights for GREV, age, majors, and UGPA, but not undergraduate institution, were significant. Of all the predictors, GREV contributed the most to the prediction of the FGPA. It is noteworthy that GREV contributed the most to the prediction of the FGPA because the study involved students who earned degrees in Verbal fields (of the 160 students, 81% or 129 of them were social sciences, behavioral sciences and humanities majors). This finding is consistent with that of Wilson (1979) in

which the researcher found that the GREV score, rather than the GREQ, tended to have higher correlation with the first year GPA in the Verbal fields.

In summary, the findings indicate that GREV, age, major, and UGPA were valid predictors of FGPA. However, the researchers were critical of the heavy weight being placed on GRE scores by some graduate schools and suggested that graduate schools/admission committees consider noncognitive variables, including "individual persistence, motivation, reasoning skills, creativity, interpersonal skills, writing skills, and prior achievement" (p. 277).

American Indian/Alaska Native students. House (1997) investigated the predictive validity of Graduate Record Examination scores for American Indian/Alaska Native students longitudinally. Data from 28 students (N = 28: 26 students in master's programs; and 2, in doctoral programs) were collected from a large public university. The predictors were GREV, GREQ, and the combined GRE scores (V+Q). The criterion measures were graduate GPA (GGPA) and degree completion status.

Data were analyzed using a correlation approach. The results showed that no significant relationships were found between the GRE scores and GGPA. However, significant relationships were found involving GREQ (r = .47, p < .05) and the combined GRE scores, V+Q (r = .40, p < .05) and students' degree completion status.

The researcher concluded that GRE scores, in some instances, significantly predict the completion of American Indian/Alaska Native students' graduate degrees. However, he added that further research is needed to assess whether the findings are similar to other institutions. Nonetheless, the GRE scores, either separate or combined, were not a valid predictor of the graduate GPA for this sample.

**Hispanic students.** Whitworth and Barrientos (1990) investigated the predictive relationships of the GRE and undergraduate GPA (UGPA) with graduate GPA (GGPA) for Hispanic and Anglo students. Predictors include GREV, GREQ, GREA, and UGPA, and the criterion measure was GGPA. Five year data were collected from a total of 952 students (N = 952: 320 Hispanic; and 632 Anglo) admitted to University of Texas at El Paso.

The data were analyzed using the regression method. The multiple correlations between a composite of V+Q+A+U and GGPA were computed, and the GGPA was regressed on the composite predictor variables. The multiple correlations were small for both Hispanics and Anglo students. The *R*s were .19 and .27, and the  $R^2$ s were .04 and .07 for Hispanics and Anglos, respectively. This indicates that approximately 4% of the GGPA can be accounted for by its linear relationship with the GREV, GREQ, GRE-A, and UGPA for Hispanics, whereas approximately 7% of the GGPA can be accounted for by the composite predictor variables for Anglos. The results showed that the magnitude of the predictive relationship of the GRE to the GGPA was a little higher for Anglos than Hispanics.

However, the standardized regression coefficients ( $\beta$ ) for GREV, GREQ, GREA, and UGPA were .00, .00, .00, .28, respectively for Hispanics. The t-tests indicated that the beta weights for all predictors were not significant. This indicates that changes in GGPA were not associated with changes in GRE scores, whereas changes in GGPA were

associated with changes in UGPA. Therefore, the GRE scores were not a valid predictor of GGPA for Hispanics.

The beta weights for Anglos were similar to those for Hispanics. The comparable  $\beta$ s were .00, .00, 00, and .25, and the UGPA was the only coefficient that was significant. In short, UGPA was found to be a predictor of the GGPA for both Hispanics and Anglos.

The researchers discussed that range restrictions in GGPA and compensatory selection practice limit the accuracy of predicting GGPA from GRE scores. But, they concluded that "GRE scores are not effective in predicting graduate school grades for either Hispanics or Anglos for this group of graduate students" (p. 131) and argued that based on the findings of the study, the practice of using GRE scores as the sole, or even a partial, basis for graduate school selection might result in discriminating against Hispanics when compared to Anglos.

The range restriction is a problem endemic to GRE validity study, and it might have contributed to the low correlations found in the study. Similarly, the effect of compensatory selection reduces the validity coefficients if the correlations among the predictors are small (Dawes, 1975; Kuncel, Hezlett, & Ones, 2001). However, sufficient relevant information was not provided in order to completely evaluate the effect of compensatory selection in that study.

### Summary of the Validity Data Based on Minority Students

Validity data based on minority students are scarce, and the validity evidence has been inconsistent. The two studies (one study based on American Indian/Alaska Native; and one based on Hispanics) found that GRE scores (both GREV and GREQ) do not tend to be associated with cumulative GPA, whereas one study (based on an African American

majority) found that only the GREV scores had significant relationship with the first year graduate GPA. In other words, it is difficult to draw any general conclusions about the predictive validity of the GRE for minority students at this time. Many more empirical research studies are needed to assess the true relationship between the GRE scores and the first year GPA, cumulative GPA, and other relevant success measures for these groups.

**Men and women.** Kaczmarek and Franco (1986) studied the predictive relationship between GRE scores and final GPA. Data from a total of 43 Caucasian students were collected in a Master's program in counseling. Eighteen students (42%) were men while 25 (58%) were women. The predictors were the GREV, GREQ, and the combined GRE (V+Q). The criterion measure was final GPA.

The data were analyzed using correlation methods. The results indicated that the GRE was not a valid predictor of final GPA for Caucasian men while the combined GRE (V+Q) and GREQ were valid predictors of final GPA for Caucasian women. The correlation of GREQ with final GPA for women was .56, while the multiple correlation of GREV and GREQ with final GPA was .52.

In discussing the results, the researchers considered range restrictions in the final GPA as the factor that reduced the validity coefficients of the GRE scores. This study is based on a small sample and one particular ethnic group (Caucasian) in one institution, thus, the findings may not be generalizable for men and women in larger populations.

# **Present Study**

As was reviewed above, extensive validity data for the GRE are available for White students. The validity evidence based on this population generally has been

consistent, indicating that the GRE scores have significant relationship with the first year GPA and cumulative GPA, where the magnitudes of the correlations were small to medium. In contrast, the GRE validity studies on minority students have been limited by small samples, and the empirical evidence about predictive validity of the GRE score from those studies has been inconsistent.

The present study aimed to investigate the validity of the GRE scores (GREV and GREQ) for predicting success in graduate school at a large racially and ethnically diverse public university in Southeast Florida. The predictive relationships of the GRE scores to success in graduate school were examined based on a population that included a large number of minority students, using regression and correlation methods. The predictors included GREV, GREQ, and UGPA. The criterion was success in graduate school, as measured by FGPA, CGPA, and degree attainment status.

### Summary

Chapter 2 reviewed the conceptual and empirical literatures relative to this GRE validity study. In the conceptual literature review, three types of theories, test theory, performance theory, and knowledge acquisition theory, were discussed to explain the rational for why the GREV and GREQ should predict the first year GPA. In empirical literature review, the validity coefficients and regression coefficients were gathered from previous GRE validity studies and reviewed by ethnic and sex groups to evaluate whether the GRE General Test scores predicted success in graduate school, as measured by first year graduate GPA, cumulative graduate GPA, and degree attainment status.

Chapter 3 delineates the research methods employed for this study. Chapter 4 presents the findings of the study, and Chapter 5 discusses the findings, implications for practice, and further research, and ends with a conclusion.

### **CHAPTER III**

### **METHOD**

This chapter includes the research design, procedures for collection of data, and the methods of data analysis. The chapter ends with a summary.

### **Research Design**

The conceptual framework for the present study is based on the criterion model (Kane, 2006), also known as the predictive validation model (Cronbach, 1970). In the predictive validation model, the validation is defined as "the process of examining the accuracy of a specific prediction or inference made from a test score" (Cronbach, 1971, p. 443). Thus, the focus of the present study was to evaluate the accuracy of the ETS prediction, that is, "GRE General Test scores are valid predictors of success in the first year of graduate school for all students" (ETS, 2009, p. 7).

Pearson's correlations, multiple linear and logistic regression, and hierarchical multiple linear and logistic regression methods were used to evaluate the relationships under study. Pearson's correlation and multiple regression methods have traditionally been used to investigate such relationships as predicting success in college or in later life based on high school grades, and college entrance examination (SAT or ACT) scores (Cohen, Cohen, West, & Aiken, 2003; Ezekiel & Fox, 1959; Stevens, 1996; Kane, 2006). The use of multiple regression and correlation methods for this study produced indices (r,  $r^2$ , R,  $R^2$ , B or  $\beta$ ) that helped determine the degree of the relationship, and evaluate the accuracy of the prediction made from a test score.

The "logistic regression is basically an extension of multiple regression in situations where the DV (dependent variable) is not a continuous or quantitative variable"

(Mertler & Vannatta, 2010, p. 289) but a dichotomous variable (Moore, & McCabe, 1998). Logistic regression method was used when the degree attainment status (attained or not attained) variable was involved in the research questions. The use of logistic regression model for this study produced an index of odds ratio (*OR*) that helped predict the probability of students' degree attainment status – either (0) degree not attained or (1) degree attained within a time limit.

Hierarchical multiple linear and logistic regression methods were used to examine possible differences in relationships between the GREV, GREQ, UGPA, as a composite predictor, and FGPA, CGPA, and degree attainment status by race/ethnicity and sex.

# Procedures

First, student records were requested by submitting a data request form to the Office of Planning and Institutional Research (OPIR) at the participating university.

Second, upon the receipt of the data, which consisted of seven Excel files, the files were converted to SPSS files. Then the data files were split into master's and doctoral degree levels. Each data file was prepared for each degree level analysis to address the research questions. Errors and duplicates were identified and deleted. All data sets that contained research variables were used unless discrepancies were found.

Third, for a preliminary analysis, frequencies and percentages, means and standard deviations were calculated for all demographic and research variables. Then, the specific research questions were answered by using Pearson's correlation, multiple linear and logistic regression, and hierarchical multiple linear and logistic regression methods. All statistical procedures were conducted using SPSS version 20.

### **Population**

The population of interest consisted of graduate students enrolled in master's and doctoral degree programs in the US. The graduate students enrolled in master's and doctoral degree programs totaled 1,698,445 (CGS, 2008). The graduate student population in the US is composed of 916,369 (72%) White, and 359,841 (28%) minority (CGS, 2008). The minority groups consist of Native American/Alaska Native, African American, Asian/Pacific Islander; and Hispanic/Latino. The sex ratio of men and women is 475,634 (38%) men and 790,146 (62%) women (CGS, 2008).

### Sample

General requirements for selecting a good sample include that samples are selected randomly, and are "representative of the target population so as not to introduce any further biases, and large enough to allow for stable and statistically significant data" (Brown, 1983, p. 122). However, sampling procedures need to be "directed toward yielding a 'good' estimate of a population characteristic" (Ackoff, 1965, p. 84).

The purpose of the present study involved examination of the predictive relationship between the GRE scores and success in graduate school by race/ethnicity and sex. This required that the sample comprises various different subpopulations "large enough to allow for stable and statistically significant data (Brown, 1983, p. 122)". Thus, a random sampling method would not be appropriate as the population is composed of a large number of White students (72% of the graduate student population are White students, CGS, 2007). Therefore, it seemed appropriate to use the purposive sampling method where a sample was selected on the basis of knowledge of a population and the purpose of the study (Babbie, 2001; Singleton, Straits, Straits, & McAllister, 1988).

With regards to the selection of the sample size, two factors have been taken into consideration. One was the requirement for the predictive validity study, and another was the nature of the criterion measures under study. First, the minimum sample size for predictive validity study is 100 subjects/cases (Anastasi, & Urbina 1997; Cronbach, 1971). Second, one of the criterion measures included degree attainment status. To study the relationships relative to this variable, an extended period of data were needed, long enough to obtain the degree attainment status. The time limits for graduation varied across the colleges and programs ranging from 6 years to 9 years. A total of 7,367 graduate student records from 2000 to 2010 were used for the study. The original data sets that contained the research variables included 5,990 students for master's programs and 1,377 for doctoral programs, respectively.

### Variables and Indicators

Variables include three predictor variables and a criterion variable with three criterion measures. The three predictor variables are GRE Verbal Reasoning Test (GREV) Scores, GRE Quantitative Reasoning Test (GREQ) Scores, and undergraduate GPA. The criterion variable is success in graduate school, as measured by first year graduate GPA (FGPA), cumulative GPA (CGPA), and degree attainment status.

The selected variables are the ones that researchers have considered the most relevant to the GRE validity study. The selected variables will be described in detail in terms of operationalization, and the quality of measures where applicable. In addition, coding procedures for subgroup variables are also included as the present study involves subgroup analyses. Then, the level of measurement, and the issue of reliability of the measurement are considered where applicable. The content and construct validity of the

measures are assumed valid as they are prerequisite to a predictive validity study (Breland, 1979; Cronbach, 1971).

### **Predictor Variables**

GRE Verbal Reasoning Test (GREV) Score: This interval level measure represents the scores of the verbal reasoning subtest of the GRE General Test. Raw scores, i.e., "Scores obtained directly from a test" (Brown, 1983, p. 38) were collected as recorded from the student record and used for data analysis. This interval level measure has the quality that enables us "to add or subtract the numbers assigned to cases in analyzing data" (Singleton, 1988, p. 109). As such, the GRE scores can be compared and stated that one number is greater or lesser than the other, and how much greater or lesser it is.

The range of the GRE Verbal scores was 200 to 800 in 10-point increments. The mean score was 457 based on the norm group that consisted of all examinees who took the test between July 1, 2005, and June 30, 2008 (ETS, 2009). Note that this was the way the GRE scores were scaled when the data used for the present study were collected. The GRE General Test and scoring scales have changed since August 2011. Also note that, although the items in the new GRE test may not be the same, the test is said to measure similar abilities and skills to those of old GRE test (ETS, 2012).

GRE Quantitative Reasoning (GREQ) Score: This interval level measure represents the scores of the quantitative reasoning subtest of the GRE General Test. Raw scores were collected as recorded from the student record and used for data analysis. While the range of scores for the GREQ is the same as the GREV, i.e., 200 to 800 in 10point increments, and the norm group was the same, the mean score was 586. Thus, it is

important to note that the GREV and GREQ should not be compared as if they are the same measure as each measure is scaled separately (ETS, 2009). The GREQ scores will be analyzed similarly to the GREV.

Undergraduate GPA (UGPA): This interval level measurement represents the average of undergraduate level course work. Raw scores utilizing the score on the 4.0 scale will be used as is in the student record. As this interval level measure can be added or subtracted, it will be used as such.

### **Criterion Variables**

Based on the related predictive validity studies, the criterion (success in graduate school) of the study is defined as consisting of the three academic variables, i.e., first year GPA (FGPA), cumulative graduate GPA (CGPA), and degree attainment status for measurement. Criterion measures consist of raw scores.

First year GPA (FGPA): This interval level measurement, on a 0 to 4 scale, represents the grade point average (GPA) of all courses students took during the first year that was calculated by the university.

Cumulative Graduate GPA (CGPA): This interval measurement, on a 0 to 4 scale, is defined differently for master's and doctoral students. Because master's students graduate in shorter period of time, three years on average (NCES, 2007), as compared to doctoral students, six years on average (NCES, 2007), master's students' CGPA is defined as GPAs cumulated over 12 months from the first enrollment in a master's degree program. For doctoral students, it is defined as GPAs cumulated over 24 months from the first enrollment in a doctoral degree program (Kuncel, et al, 2000).

Degree attainment status: This nominal level measure will be coded 0 for degree not attained, and 1 for degree attained. The variable, degree attainment status, had two levels, 'degree attained' and 'degree not attained'. Students who attained degrees within the time limit set by the university were included in the category of 'degree attained', and students who either did not attain the degrees within the time limit or dropped out of the program were included in the category of 'degree not attained'. The time limits were 6 years for master's students, and 9 years for doctoral students (FIU, 2011).

### **Subgroup Variables**

Sex: This nominal level measure was coded 0 for female and 1 for male. Numerals are assigned to the categories for the convenience of the study in analyzing data.

Race/Ethnicity: This nominal level measure was categorized as African American, Asian American, Hispanic, Native American, White, International students, and NRs whose race/ethnicity were not reported. Each race/ethnic group was coded 0 and 1, with the particular group of interest coded as 1, and the others 0.

### **Data Collection**

The student data of interest included admission status, nationality, race/ethnicity, age, sex, college, degree level, department, entry term, exit term, GRE verbal score, GRE quantitative score, undergraduate GPA, first year graduate GPA, cumulative graduate GPA, and degree attainment status. Student records were obtained by submitting a data request form to the university's Office of Planning and Institutional Research (OPIR). The data were collected from all individuals who were enrolled in master's and doctoral programs during the 10 academic years, from 2000 to 2010.

### **Data Analysis**

Data were analyzed separately for two degree levels, masters and doctoral, for two reasons: First, two of the research variables (Cumulative GPA and degree attainment status) were defined differently for master's and doctoral degree levels. Second, the race/ethnicity composition of the students of the two degree levels was different: Hispanic students were the largest group in master's degree programs while White students were the largest group in doctoral degree programs. The sample was also divided by subgroups to answer the research questions as appropriate. The subgroups consisted of two sex groups (men and women) and seven racial/ethnic/other groups: African American, Asian American, Hispanic, Native American, White, International students, and Not-Reported groups. The Not-Reported (NR) group represents students whose race/ethnic identity was not reported.

Pearson's correlations, multiple linear and logistic regression, and hierarchical multiple linear and logistic regression methods were used to answer the four research questions. Statistical methods and procedures to answer the specific research questions are explained as follows:

### **Research Question One**

To what degree do GRE scores taken singly (individual GREV and GREQ), and in total, sum of GREV and GREQ (GRET), predict success in graduate school as measured by FGPA, CGPA, and degree attainment status?

Pearson's correlation analyses were used to determine the degree of the relationship of the GRE scores with FGPA, CGPA, and degree attainment status for master's and doctoral degree level analyses. The degree of the relationship was expressed

as a correlation coefficient (*r*). The correlations of .10, .30, and .50, regardless of sign, were interpreted as small, medium, and large correlations, respectively (Cohen, 1988). If the correlations were significant and positive, the relationship was interpreted as higher FGPAs (CGPAs, degree attainment) were associated with higher GREV (GREQ, GRET).

### **Research Question Two**

To what degree does undergraduate GPA (UGPA) predict success in graduate school, as measured by first year graduate GPA, cumulative GPA, and degree attainment status?

The question was similar to the Question 1, but with different predictor. Pearson correlation analyses were used to determine the degree of the relationship between UGPA and the three success measures (FGPA, CGPA, and degree attainment status) for master's and doctoral degree level analyses. If the correlations were significant and positive, the relationship was interpreted as high FGPAs (high CGAP; degree attainment) were associated with high UGPAs.

### **Research Question Three**

To what degree do GREV, GREQ, and UGPA, as a composite predictor, predict success in graduate school, as measured by FGPA, CGPA, and degree attainment status?

Multiple linear regression analyses were performed on FGPA (CGPA) with GREV, GREQ, and UGPA, as a composite predictor, to determine the degree of relationship between FGPA (CGPA), and the composite predictor for master's and doctoral degree level analyses. The degree of the relationship was examined in terms of multiple correlation coefficient ( $R^2$ ), regression coefficients (B), and the standardized regression coefficients ( $\beta$ ). Each model and the indices were tested for significance. The

 $R^2$  is the indicator that determines whether the composite predictor (GREV, GREQ, and UGPA) have a relationship with a criterion (FGPA or CGPA). If the  $R^2$  were significant  $(R^2 > 0)$ , the relationship was interpreted as  $R^2$  percent of the variability in the criterion measure (FGPA or CGPA) was explained by the composite predictor. The coefficients, *B* and  $\beta$ , were used in evaluating the relative importance of each predictor in predicting the criterion measure (FGPA or CGPA).

Multiple logistic regression analyses were performed on degree attainment status with GREV, GREQ, and UGPA, as a composite predictor, to determine the relationship between degree attainment status and the composite predictor for master's and doctoral degree level analyses. The degree of the relationship was examined in terms of two Rsquared measures (Cox & Snell R square and Nagelkerke R square), odds ratios (*OR*s), and 95% confidence intervals for *OR*s. Each model and the odds ratios were tested for significance. The odds ratio indicates a percent change in the odds of criterion measure (degree attainment) for one point increase in the corresponding predictor (e.g., GREV), holding the other predictors (e.g., GREQ, and UGPA) constant. If the model was significant, the two  $R^2$  (Cox & Snell R square and Nagelkerke R square) were interpreted as  $R^2$  percent of the variability in degree attainment explained by the composite predictor.

### **Research Question Four**

Do the relationships between GREV, GREQ, and UGPA, as a composite predictor, and success in graduate school, as measured by FGPA, CGPA, and degree attainment status, differ by race/ethnicity and sex?

Hierarchical multiple linear regression analyses on FGPA and CGPA and hierarchical multiple logistic regression analyses on degree attainment status were

performed to measure the differences by race/ethnicity and sex at both masters and doctoral level analyses. In all models, the composite predictor included GREV, GREQ, UGPA, male (this is the label used for statistical analyses to indicate sex), race/ethnicity, and interactions between each of the composite predictor and male, and interactions between each of the composite predictor and each race/ethnicity. The White student group was used as the reference group in all of the regression models. The criterion variables were FGPA, CGPA, and degree attainment status.

To test for the significance of the interaction of race/ethnicity by the composite predictor, two regression models were fitted and tested for differences. In the first step, a multiple linear regression model (or a multiple logistic regression model) was constructed with main effects of the composite predictor, male, race/ethnicity, and cross products of male with each composite predictor. In the second step, cross products of race/ethnicity with each composite predictor were added. The  $R^2$  change (or the difference in the chi-square statistics) between the two models tested the significance of the interaction of race/ethnicity by the composite predictor.

Similarly, to test for the significance of the interaction of male by the composite predictor, in the first step, a multiple linear regression model (or a multiple logistic regression model) was constructed with main effects of the composite predictor, male, race/ethnicity, and cross products of race/ethnicity with each of the composite predictor, and in the second step, cross products of male with each of the composite predictor were added. The  $R^2$  change (or the difference in the chi-square statistics) between the two models tested the significance of the interaction of male by the composite predictor.

Differences in predictive relationships between FGPA and CGPA and the composite predictor among race/ethnicity and sex groups were examined in terms of significance (p - value) of the regression equation, magnitude of  $R^2$  change, and regression coefficients (B, and  $\beta$ ) of the individual predictors. Differences in predictive relationships between degree attainment status and the composite predictor among race/ethnicity and sex groups were examined in terms of significance (p - value) of the logistic regression equation, pseudo  $R^2$  (Cox and Snell  $R^2$ , Nagelkerke  $R^2$ ), odds ratios (*ORs*) and 95% confidence intervals for *ORs*.

### **Summary**

Chapter 3 discussed the research design delineating the population, the sample, measurement, the data collection, and the data analysis of the study. The sample was purposely selected to evaluate the predictive relationships between the GREV, GREQ and UGPA and success in graduate school, as measured by FGPA, CGPA, and degree attainment, for the population that includes a large number of minority students.

The data were collected from the archives of the university's Office of Planning and Institutional Research. Both raw scores and dummy codes were used in data analyses. The data were analyzed using Pearson's correlations, multiple linear and logistic regression, and hierarchical multiple linear and logistic regression methods. Chapter 4 presents the findings of the study. Chapter 5 discusses the findings, implications for practice, and further research, and ends with a conclusion.

### **CHAPTER IV**

### RESULTS

This chapter presents the results of the research study that examined the validity of the GRE General Test Scores (GREV and GREQ) and undergraduate GPA (UGPA) for predicting success in graduate school, as measured by three success measures: first year graduate GPA (FGPA), cumulative graduate GPA (CGPA), and degree attainment status. The predictive validity was examined using Pearson's correlations, multiple linear (and logistic) regression, and hierarchical multiple linear (and logistic) regression methods. Frequencies and percentages, means and standard deviations were calculated for all demographic and research variables. All statistical analyses were performed using the SPSS version 20. The probability level was set at .05,  $\alpha = .05$ , to determine the significance of the relationship.

The research questions that guided the study were as follows:

- To what degree do GRE scores taken singly (individual GRE Verbal Reasoning Test score and GRE Quantitative Reasoning Test score), and in total, sum of GREV and GREQ (GRET), predict success in graduate school, as measured by FGPA, CGPA, and degree attainment status?
  - a. To what degree does GREV predict success in graduate school?
  - b. To what degree does GREQ predict success in graduate school?
  - c. To what degree does GRET predict success in graduate school?
- 2. To what degree does UGPA predict success in graduate school, as measured FGPA, CGPA, and degree attainment status?

- 3. To what degree do GREV, GREQ, and UGPA, as a composite predictor, predict success in graduate school, as measured by FGPA, CGPA, and degree attainment status?
- 4. Do the relationships between GREV, GREQ and UGPA, as a composite predictor, and success in graduate school, as measured by FGPA, CGPA, and degree attainment status, differ by race/ethnicity and sex?

Data were analyzed separately for two degree levels, master's and doctoral, for two reasons: First, two of the research variables (cumulative GPA and degree attainment status) were defined differently for the two degree levels. Second, the ethnic composition of the students of the two degree levels were different (Hispanics were the largest group in master's programs while White students were the largest group in doctoral programs). The sample was also divided by subgroups to answer research questions as appropriate. The subgroup consisted of two sex groups (men and women), and seven racial/ethnic/other groups: African American, Asian American, Hispanic, Native American, White, International students, and Not-Reported groups. The Not-Reported (NR) group represents students whose race/ethnicities were not reported.

### Sample

Master's and doctoral student records at the participating university were used for the study. The original data sets that contain the research variables included 5,990 students for master's programs and 1,377 for doctoral programs, respectively. However, different sample sizes (N) were used for each research question, as the variables involved in each research question differed from each other, and the data sets that contained the research variables were also different.

# **Demographic Variables**

Note that racial/ethnic categories in this research study were applicable to U.S. citizens only. Students from foreign countries were categorized as international students in the research study.

As shown in Table 1, the largest percentage of race/ethnicity for master's students was Hispanics (39.5%), followed in order by White (28.3%), international students (13.8%), African Americans (13.2%), Asian Americans (3.8%), Not Reported (NR) (1.2%), and Native Americans (0.2%). For doctoral students, the largest percentage of race/ethnicity was White (31.4%), followed in order by international students (29.4%), Hispanics (23.7%), African Americans (9.8%), Asians (3.4%), NR (1.7%), and Native Americans (0.4%).

Table 1

		Master's	s Students	Doctoral Students		
Characteristic	Variable	f	%	f	%	
Sex	Women	3,722	62.1	710	51.6	
	Men	2,268	37.9	667	48.4	
	Total	5,990	100.0	1,377	100.0	
Race/Ethnicity	African American	789	13.2	135	9.8	
	Asian American	229	3.8	47 <sup>a</sup>	3.4	
	Hispanic	2,369	39.5	327	23.7	
	International students	829	13.8	405	29.4	
	Native American	10 <sup>a</sup>	0.2	6 <sup>a</sup>	0.4	
	NR	70	1.2	24 <sup>a</sup>	1.7	
	White	1,694	28.3	433	31.4	
	Total	5,990	100.0	1,377	100.0	

Summary of Demographic Variable

*Note.* White is an abbreviation of White, non-Hispanic. NR is a group whose racial/ethnic identity was not reported. <sup>a</sup> For both masters and doctoral degree level analyses, Native Americans were treated as missing in Research Question Four due to small sample sizes of all research variables. For doctoral degree level analysis only, Asian Americans were also treated as missing in Research Question Four due to a small sample size. In addition, NRs at doctoral degree level were excluded in Research Question Four due to missing data for CGPA and degree attainment variables.

For master's students, approximately 62% were women, and 38% were men. For doctoral students, approximately 52% were women, and 48% were men.

As shown in Table 2 and 3, the average age for master's students was 29 (SD = 7.89), and the ages ranged from 16 to 79 years old. For doctoral students, the average age was 31(SD = 8.96) and the ages ranged from 20 to 70 years old.

### **Research Variables**

**GRE scores.** As shown in Tables 2, the mean GREV score for master's students was 433.01 (SD = 100.99), and the scores ranged from 200 to 800 (N = 5,990). As shown in Tables 3, the mean GREV score for doctoral students was 491.06 (SD = 107.50), and the scores ranged from 250 to 800 (N = 1,377). For master's students, the mean GREQ score was 530.65 (SD = 138.11), and the scores ranged from 200 to 800. For doctoral students, the mean GREQ score was 628.52 (SD = 125.78), and the scores ranged from 200 to 800. For master's students, the mean GREQ score was 963.54 (SD = 194.26), and the scores ranged from 440 to 1,590. For doctoral students, the mean GRET score was 1,119.49 (SD = 176.04), and the scores ranged from 480 to 1,590.

Undergraduate GPA. For master's students, the mean UGPA was 3.33 (SD = 0.42), and the grades ranged from 0.62 to 4.00 (N = 5,778). For doctoral students, the mean UGPA was 3.52 (SD = 0.38), and the grades ranged from 1.76 to 4.00 (N = 1,325).

First year graduate GPA. For master's students, the mean FGPA was 3.50 (*SD* = 0.54), and the grades ranged from 0.00 to 4.00 (N = 5,948). For doctoral students, the mean FGPA was 3.65 (*SD* = 0.43), and the grades ranged from 0.00 to 4.00 (N = 1,377).

# Table 2

Means, Standard Deviations, and Intercorrelations among Variables for Master's Students for Research Question 1

			1.	2.	3.	4.	5.	6.	7.	8.	9.
Variables	M	SD	GREV	GREQ	GRET	UGPA	FGPA	CGPA	Deg6	Deg3	Age
1. GREV	433.01	100.98									
2. GREQ	530.65	138.11	0.30**								
3. GRET	963.54	194.26	0.74**	0.87**							
4. UGPA	3.33	0.42	0.07*	-0.01	0.04**						
5. FGPA	3.50	0.54	0.15***	0.05***	0.12***	0.21***					
6. CGPA	3.41	0.76	0.10***	0.06***	0.09***	0.17***	0.49**				
7. Deg6	84.1 <sup>a</sup>		-0.10***	-0.00	-0.05***	0.06***	0.18**	0.10**			
8. Deg3	73.3 <sup>b</sup>		-0.11***	-0.09***	-0.12***	0.09***	0.22**	0.11**	1.00**		
9. Age	29.0 <sup>c</sup>	7.89	0.07**	-0.21**	-0.12**	-0.08**	0.01	-0.02	-0.06**	-0.04**	

*Note.* N = 4,833 to 5,990. GREV is GRE Verbal Reasoning Test scores, GREQ is GRE Quantitative Reasoning Test score, GRET is the total of GREV and GREQ scores, UGPA is undergraduate GPA, FGPA is first year graduate GPA, CGPA is cumulative GPA, Deg6 is degree attainment status within 6 year time limit, and Deg3 is degree attainment status within 3 years. The numbers on lines with 7 and 8 in the column *M* are the percent of master' students who have attained degrees within the specified times: <sup>a</sup> about 84% or 4,357 out of 5,181(from 2000 to 2010 academic years) master's students graduated within 6 years; <sup>b</sup> about 73% (3,829 out of 5,222) master's students graduated within 3 years while 27% or 1,393 master's students have either not graduated within 3 years or dropped out of the program; <sup>c</sup> the mean age of master's students was 29. The master's students' ages ranged from 16 to 79 years old.

\**p* < .05. \*\**p* < .01. \*\*\* *p* < .001.

# Table 3

Means, Standard Deviations, and Intercorrelations among Variables for Doctoral Students for Research Question 1

			1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
Variables	M	SD	GREV	GREQ	GRET	UGPA	FGPA	CGPA	Deg9	Deg8	Deg6	Age
1. GREV	491.06	107.50										
2. GREQ	628.52	125.78	0.13**									
3. GRET	1,119.49	176.04	0.71**	0.80**								
4. UGPA	3.52	0.38	0.10**	-0.03	0.04							
5. FGPA	3.65	0.43	0.07**	0.03	0.07*	0.19***						
6. CGPA	3.63	0.49	0.06	0.05	0.07*	0.13***	0.58**					
7. Deg9	75.8 <sup>a</sup>		0.05	0.14**	0.13**	0.12*	0.25**	0.18**				
8. Deg8	67.6 <sup>b</sup>		0.06	0.16***	0.15**	0.09	0.27**	0.21**	1.00**			
9. Deg6	50.1 <sup>c</sup>		0.00	0.12**	0.09*	0.05	0.25**	0.18**	1.00**	1.00**		
10. Age	31.0 <sup>d</sup>	8.96	0.02	-0.43**	-0.30**	-0.03	-0.01	-0.01	-0.25**	-0.25**	-0.22**	

*Note*.  $N = 400 \sim 1,377$ . GREV is GRE Verbal Reasoning Test scores, GREQ is GRE Quantitative Reasoning Test score, GRET is the total of GREV and GREQ scores, and UGPA is undergraduate GPA, FGPA is first year graduate GPA, CGPA is cumulative graduate GPA, Deg9 is degree attainment status within a 9-year time limit, Deg8 is degree attainment status within eight years, Deg6 is degree attainment status within 6 years. The numbers on lines with 7, 8, and 9 in the column *M* are the percent of doctoral students who have attained degrees within the specified times: <sup>a</sup> about 76% or 338 out of 446 doctoral students have graduated within 9 years. About 24% or 108 doctoral students have either not graduated within 9 years or dropped out of the program; <sup>b</sup> about 68% or 326 out of 482 doctoral students (from 2000 to 2010 academic years) have graduated within 8 years; about 32% or 156 students have either not graduated within 8 years or dropped out of the program; <sup>c</sup> about 50% or 269 out of 537 doctoral students have graduated within 6 years; about 50% or 268 students have either not graduated within 6 years or dropped out of the program; <sup>d</sup> the mean age of doctoral students was 31. The doctoral students' ages ranged from 20 to 79 years old.

\*p < .05. \*\* p < .01. \*\*\* p < .001.

**Cumulative graduate GPA.** For master's students, the mean CGPA was 3.41 (SD = 0.76), and the grades ranged from 0.00 to 4.00 (N = 4,833). For doctoral students, the mean CGPA was 3.63 (SD = 0.49), and the grades ranged from 0.00 to 4.00 (N = 1,059).

**Degree attainment Status.** For master's students, approximately 84% (4,357 of 5,181) attained their degrees within a 6-year time limit. Approximately 16% (824) of the students either did not attain their degrees within six years or dropped out of the programs. The mean number of years to attain a master's degree was approximately three years (M = 36.22 months, SD = 30.84 months, range = 0 to 269 months).

For doctoral students, approximately 76% (338 of 446) of doctoral students attained their degrees within a 9-year time limit. Approximately 24% (108) of the students either did not attain their degrees within nine years or dropped out of the program. The mean number of years to attain a doctoral degree was approximately five years (M = 60.3 months, SD = 27.16 months, range = 3 to 281 months).

### **Research Question One: Correlational Analyses**

The first research question was to evaluate the predictive relationship between the GRE scores (individual GREV, GREQ, and the sum of GREV and GREQ) and success in graduate school, as measured by three success measures, FGPA, CGPA, and degree attainment status. The results are presented in Tables 2 and 3.

### **Relationship between the Three Success Measures and the GRE Scores**

Pearson's correlation analyses were used to determine the relationship between the GRE scores and the three success measures. The relationships were examined in terms of significance (p – value) and magnitude of the correlation coefficients (r). Relationship between the three success measures and master's students' GREV scores. The relationship between master's students' FGPA and GREV was significant (r = .15, p < .001) and positive. Higher FGPAs were associated with higher GREV scores. The relationship between master's students' CGPA and GREV was statistically significant (r = .10, p < .001) and positive. Higher CGPAs were associated with higher GREV scores. The relationship between students' degree attainment within a 6-year time limit and GREV was significant, but negative (r = .10, p < .001). Degree attainment within a 6-year time limit was associated with lower GREV scores.

Results indicate that master's students' FGPAs and CGPAs were positively associated with GREV scores. Higher FGPAs and higher CGPAs were associated with higher GREV scores. However, master's students' degree attainment within a 6-year time limit was negatively associated with GREV scores. Master's students' degree attainment within a 6-year time limit was associated with lower GREV scores.

Relationship between the three success measures and doctoral students' GREV scores. The relationship between doctoral students' FGPA and GREV was significant (r = .07, p < .01) and positive. Higher FGPAs were associated with higher GREV scores. The relationship between doctoral students' CGPA and GREV was not significant (r = .06, p = .053). Doctoral students' CGPAs were not associated with GREV scores. The relationship between doctoral students' degree attainment within a 9-year time limit and GREV was not significant (r = .05, p = .317). Thus, the results indicate that higher FGPAs were associated only with higher GREV scores for doctoral students.

In summary, for both master's and doctoral students, higher FGPAs were associated with higher GREV scores (master's r = .15, doctoral r = .07). For master's students only, higher CGPAs were associated with higher GREV scores (r = .10) while higher GREV scores were negatively associated to degree attainment in 6 years (r = .10).

# Relationship between the three success measures and master's students' GREQ scores. The relationship between master's students' FGPA and GREQ was significant (r = .05, p < .001) and positive. Higher FGPAs were associated with higher GREQ scores. The relationship between master's students' CGPA and GREQ was significant (r = .06, p < .001) and positive. Higher CGPAs were associated with higher GREQ scores. The relationship between students' degree attainment within a 6-year time limit and GREQ was not significant (r = .004, p = .754). Master's students' degree attainment within a 6-year time limit was not associated with GREQ scores. Results indicate that higher FGPAs and higher CGPAs were associated with higher GREQ scores for master's students.

### Relationship between the three success measures and doctoral students'

**GREQ scores.** Neither the relationships between doctoral students' FGPA and GREQ (r = .03, p = .207) nor CGPA and GREQ (r = .05, p = .122) was statistically significant. Doctoral students' FGPAs and CGPAs were not significantly associated with GREQ scores. However, the relationship between doctoral students' degree attainment within a 9-year time limit and GREQ was significant and positive (r = .14, p < .01). Doctoral students' degree attainment within a 9-year time limit within a 9-year time limit was associated with higher GREQ scores. Results indicate that doctoral students' degree attainment within a 9-year time limit was associated with higher GREQ scores.

In summary, for master's students, higher FGPAs (r = .05) and higher CGPAs (r = .06) were associated with higher GREQ scores. For doctoral students, degree attainment within a 9-year time limit was associated with higher GREQ scores (r = .14).

# Relationship between the three success measures and master's students' GRET scores. The relationship between master's students' FGPA and GRET was significant (r = .12, p < .001) and positive. Higher FGPAs were associated with higher GRET scores. The relationship between master's students CGPA and GRET was significant (r = .09, p < .001) and positive. Higher CGPAs were associated with higher GRET scores. The relationship between master's students' degree attainment within a 6-year time limit and GRET was significant but negative (r = ..05, p < .001). Results indicate that master's students' higher FGPAs and higher CGPAs were associated with higher FGPAs were associated with higher FGPAs and higher CGPAs were associated with higher SRET scores.

### Relationship between the three success measures and doctoral students'

**GRET scores.** The relationship between doctoral students' FGPA and GRET was significant (r = .07, p < .05) and positive. Higher FGPAs were associated with higher GRET scores. The relationship between doctoral students' CGPA and GRET was significant (r = .07, p < .05) and positive. Higher CGPAs were associated with higher GRET scores. The relationship between doctoral students' degree attainment within a 9-year time limit and GRET was also significant and positive (r = .13, p < .01). Results indicate that all three success measures, FGPA, CGPA, and degree attainment within time limit, were positively associated with GRET scores for doctoral students.

In summary, for masters students, higher FGPAs (r = .12), and higher CGPAs (r = .09) were associated with higher GRET scores. For doctoral students, all three success

measures, FGPA (r = .07), CGPA (r = .07), and degree attainment within a 9-year time limit (r = .13), were associated positively with GRET scores. Higher FGPAs, higher CGPAs, and degree attainment within a 9-year time limit were associated with higher GRET scores.

Overall, the correlations between the three success measures and the GRE scores (GREV, GREQ, and GRET) varied by degree level. For master's students, FGPAs and CGPAs were positively associated with all three GRE scores (individual GREV, GREQ, and in total, GRET). That is, higher FGPAs and higher CGPAs were associated with higher GREV, higher GREQ, and higher GRET scores. But none of the GRE scores were positively associated with master's students' degree attainment within a 6-year time limit.

For doctoral students, all three success measures, FGPA (r = .07), CGPA (r = .07), and degree attainment within a 9-year time limit (r = .13) were positively associated with GRET scores, although the other relationships between the three success measures and individual GRE scores varied: doctoral students' FGPAs were associated with GREV scores, but not with GREQ scores; doctoral students' CGPAs were associated neither with GREV nor with GREQ scores; doctoral students' degree attainment within a 9-year time limit (r = .13) was associated positively with GREQ scores, but not with GREV scores.

### **Research Question Two: Correlational Analyses**

The second research question was to evaluate the predictive relationship between undergraduate GPA (UGPA) and success in graduate school, as measured by the three success measures (FGPA, CGPA, and students' degree attainment status). The results are presented in Tables 2 and 3.
#### **Relationship between the Three Success Measures and UGPA**

Research Question Two is similar to Research Question One in examining the relationships that involve the same criterion variable (and success measures), but with a different predictor variable. Similar analyses were performed with UGPA as the predictor. Pearson's correlation analyses were used to answer the research question. The predictive relationships were examined in terms of significance (p – value) and the magnitude of the correlation coefficients (r).

Relationship between the three success measures and master's students' UGPA. The relationship between master's students' FGPA and UGPA was significant (r = .21, p < .001) and positive. Higher FGPAs were associated with higher UGPAs. The relationship between master's students' CGPA and UGPA was significant (r = .17, p < .001) and positive. Higher CGPAs were associated with higher UGPAs. The relationship between master's students degree attainment within a 6-year time limit and UGPA was also significant (r = .06, p < .001) and positive. Master's students' degree attainment within a 6-year time limit was associated with higher UGPAs.

Results indicate that all three success measures, FGPA, CGPA, and degree attainment within time limit, were positively associated with UGPA for master's students.

#### Relationship between the three success measures and doctoral students'

**UGPA.** The relationship between doctoral students' FGPA and UGPA was significant (r = .19, p < .001) and positive. Higher FGPAs were associated with higher UGPAs. The relationship between doctoral students' CGPA and UGPA was significant (r = .13, p < .001) and positive. Higher CGPAs were associated with higher UGPAs. The relationship

between doctoral students' degree attainment within a 9-year time limit and UGPA was significant (r = .12, p = .02) and positive. Students' degree attainment within a 9-year time limit was associated with higher UGPAs.

Results indicate that all three success measures, FGPA, CGPA, and degree attainment within a 9-year time limit, were positively associated with UGPA for doctoral students.

In summary, for both master's and doctoral students, the correlations between the three success measures and UGPAs were all significant and positive. Higher FGPAs (master's r = .21, doctoral r = .19), higher CGPAs (master's r = .17, doctoral r = .13), and students' degree attainment (master's r = .06, doctoral r = .12) were associated with higher UGPAs.

### **Research Question Three: Multiple Linear and Logistic Regression Analyses**

The third research question was to evaluate the predictive relationship between the composite predictor (GREV, GREQ, and UGPA) and success in graduate school, as measured by three success measures (FGPA, CGPA, and degree attainment status). See Tables 4 - 9 for presentation of findings.

#### Relationship between the Three Success Measures and the Composite Predictor

Two multiple linear regression analyses (on FGPA and CGPA) and a multiple logistic regression analysis (on degree attainment status) were performed to answer the research question. The predictive relationships based on the two multiple linear regression analyses were examined in terms of significance (p – value) of the regression equation, the magnitude of the multiple correlation coefficients ( $R^2$ , and adjusted  $R^2$ ), and regression coefficients (B, and  $\beta$ ) of the three individual predictors. The predictive

relationships based on the multiple logistic regression analyses were examined in terms of the significance (p – value) of the regression equation, the pseudo  $R^2$  (Cox and Snell  $R^2$ , Nagelkerke  $R^2$ ), logistic regression coefficients (B), odds ratios (OR) and 95% confidence intervals of odds ratios of the three individual predictors.

# Relationship between FGPA and the composite predictor.

## Relationship between master's students' FGPA and the composite predictor.

The regression equation of master's students' FGPA on the composite predictor was significant,  $R^2 = .063$ , adjusted  $R^2 = .063$ , F(3, 5, 735) = 129.41, p < .001. As shown in Table 4, about 6% of the variability in the master's students' FGPA was explained by the linear combination of the three predictors (GREV, GREQ, and UGPA).

The regression coefficients of the three individual predictors showed that two predictors, GREV and UGPA, were significant in predicting FGPA, with UGPA a stronger predictor of FGPA, B = .26 ( $\beta = .20$ ), p < .001, than GREV, B = .001 ( $\beta = .13$ ), p < .001. GREQ was not a significant predictor in this model, B = .0001 ( $\beta = .02$ ), p = .239. Results show that master students' higher FGPAs were associated with higher GREV and higher UGPA scores.

Table 4

Predictor В SE β р Constant 2.301 .064 <.001 --GREV\*\*\* .129 <.001 .00069 .000 .239 GREQ .00006 .000 .016 UGPA\*\*\* 26106 .016 .203 <.001

*Regression Analysis of Master's Students' FGPA on the Composite Predictor for Research Question 3* 

*Note*. GREV is GRE Verbal Reasoning Test scores. GREQ is GRE Quantitative Reasoning Test score. UGPA is undergraduate GPA. Model  $R^2 = .063$ , F(3, 5735) = 129.41, p < .001. \*\*\* p < .001.

# Relationship between doctoral students' FGPA and the composite predictor.

The regression equation of doctoral students' FGPA on the three composite predictor was significant,  $R^2 = .042$ , adjusted  $R^2 = .039$ , F(3, 1321) = 19.07, p < .001. As shown in Table 5, about 4% of the variability in the doctoral students' FGPA was explained by the linear combination of the composite predictor (GREV, GREQ, and UGPA).

The regression coefficients of the three individual predictors showed that UGPA was significant in predicting FGPA, B = .215 ( $\beta = .19$ ), p < .001. GREV, B = .00021 ( $\beta = .05$ ), p = .059, and GREQ scores were not a significant predictor in this model, B = .00012 ( $\beta = .04$ ), p = .192. Results show that doctoral students' higher FGPAs were associated with higher UGPAs.

Table 5

*Regression Analysis of Doctoral Students' FGPA on the Composite Predictor for Research Question 3* 

Predictor	В	SE	β	Р	
Constant	2.708	.130		< .001	
GREV	.00021	.000	.052	.059	
GREQ	.00012	.000	.036	.192	
UGPA***	.215	.031	.189	< .001	

*Note.* GREV is GRE Verbal Reasoning Test scores. GREQ is GRE Quantitative Reasoning Test score. UGPA is undergraduate GPA. Model  $R^2 = .042$ , F(3, 1321) = 19.07, p < .001. \*\*\* p < .001.

In summary, for both master's and doctoral students, the regression equations of FGPA on the composite predictor were significant (masters,  $R^2 = .063$ ; doctoral,  $R^2 = .042$ ) at *ps* < .001. About 4% (doctoral) and 6% (master's) of the variability in students' FGPAs were explained by the linear combination of the composite predictor (GREV,

GREQ, and UGPA). In addition, among individual predictors, for both master's and doctoral students, higher FGPAs were associated with higher UGPAs (master's  $\beta$  = .20, doctoral  $\beta$  = .19).

# Relationship between CGPA and the composite predictor. The second

multiple linear regression analysis was performed to evaluate the predictive relationship between CGPA and the composite predictor (GREV, GREQ, and UGPA).

# Relationship between master's students' CGPA and the composite predictor.

The regression equation of master's students' CGPA on the composite predictor was significant,  $R^2 = .036$ , adjusted  $R^2 = .035$ , F(3, 4651) = 57.34, p < .001. As shown in Table 6, about 4% of the variability in the master's students' CGPA was explained by the linear combination of the composite predictor (GREV, GREQ, and UGPA).

Table 6

*Regression Analysis of Master's Students' CGPA on the Composite Predictor for Research Question 3* 

Predictor	В	SE	β	Р	
Constant	2.077	.103		< .001	
GREV***	.00053	.000	.071	< .001	
GREQ*	.00021	.000	.038	.013	
UGPA***	.29541	.026	.162	< .001	

*Note.* GREV is GRE Verbal Reasoning Test scores. GREQ is GRE Quantitative Reasoning Test score. UGPA is undergraduate GPA. Model  $R^2 = .036$ , F(3, 4651) = 57.34, p < .001.\*p < .05.\*\*\*p < .001.

The regression coefficients of the three individual predictors showed that all three predictors were significant in predicting master's students' CGPA. UGPA was the strongest predictor of CGPA, B = .30 ( $\beta = .16$ ), p < .001, followed in order by GREV, B = .001 ( $\beta = .07$ ), p < .001, and GREQ, B = .00021 ( $\beta = .04$ ), p = .013. Results show that

master's students' higher CGPAs were associated with higher UGPAs, higher GREV, and higher GREQ scores.

# Relationship between doctoral students' CGPA and the composite predictor.

The regression equation of doctoral students' CGPA on the composite predictor was significant,  $R^2 = .022$ , adjusted  $R^2 = .019$ , F(3, 1009) = 7.40, p < .001. As shown in Table 7, about 2% of the variability in the doctoral students' CGPA was explained by the linear combination of the composite predictor (GREV, GREQ, and UGPA).

The regression coefficients of the three individual predictors showed that UGPA was significant in predicting doctoral students' CGPA, B = .16 ( $\beta = .13$ ), p < .001. But, GREV and GREQ scores were not significant predictors in this model. For GREV,

 $B = .00016 \ (\beta = .04), p = .251 \ \text{and for GREQ}, B = .00020 \ (\beta = .05), p = .109.$ 

In summary, for both master's and doctoral students, the regression equations of CGPA on the composite predictor were significant (masters,  $R^2 = .036$ ; doctoral,  $R^2 = .022$ ) at *ps* < .001. About 2% (doctoral) to 4% (master's) of the variability in students' CGPAs were explained by the linear combination of the composite predictor.

Table 7

Р Predictor В SE ß <.001 Constant 2.853 .166 --GREV .00016 .000 .037 .251 GREQ .00020 .000 .051 .109 UGPA\*\*\* .16310 .040 .128 < .001

*Regression Analysis of Doctoral Students' CGPA on the Composite Predictor for Research Question 3* 

*Note.* GREV is GRE Verbal Reasoning Test scores. GREQ is GRE Quantitative Reasoning Test score. UGPA is undergraduate GPA. Model  $R^2 = .022$ , F(3, 1009) = 7.40, p < .001. \*\*\*p < .001. In addition, for both master's and doctoral students, higher CGPAs were associated with higher UGPAs (master's  $\beta = .16$ , doctoral  $\beta = .13$ ).

#### Relationship between degree attainment status and the composite predictor.

Multiple logistic regression analyses were performed to evaluate the predictive relationship between degree attainment status within the time limit (six years for master's students; and nine years for doctoral students) and the composite predictor (GREV, GREQ, and undergraduate GPA).

#### Relationship between master's students' degree attainment status and the

*composite predictor.* The logistic regression equation of master's students' degree attainment status within six years on the composite predictor was significant,

 $\chi^2$  (3, N = 5,021) = 74.13, p < .001, Pseudo  $R^2$  = .015 to .025. As shown in Table 8,

between 1.5% and 2.5% of the variability in the master's students' degree attainment

within a 6-year time limit was explained by the composite predictor.

Table 8

Logistic Regression of Master's Students' Degree Attainment Status within a 6-Year Time Limit on the Composite Predictor for Research Question 3

			Wald's						
Predictor	В	SE	χ2	Df	р	OR	959	% CI	
GREV***	003	.000	54.51	1	<.001	.997	.996	.998	
GREQ*	.001	.000	5.73	1	.017	1.001	1.000	1.001	
UGPA***	.445	.088	25.52	1	<. 001	1.561	1.313	1.855	
Constant	1.111	.346	10.34	1	.001	3.039			

*Note.* GREV is GRE Verbal Reasoning Test scores. GREQ is GRE Quantitative Reasoning Test score. UGPA is undergraduate GPA. Model  $\chi^2$  (3, N = 5021) = 74.13, p < .001. Cox and Snell  $R^2$  = .015,Nagelkerke  $R^2$  = .025. \*p < .05. \*\*\* p < .001.

The coefficients of the three individual predictors showed that two predictors,

UGPA and GREQ were significant (UGPA, p < .001, GREQ, p = .017) and positive

(UGPA, B = .45, GREQ, B = .001) in predicting the master's students' degree attainment within six years. UGPA was a stronger predictor of degree attainment, OR = 1.56, 95% CI (1.31, 1.86), than GREQ, OR = 1.001, 95% CI (1.000, 1.001). The coefficient of GREV was significant, p < .001, but negative (B = -.003) in this model, OR = .997, 95% CI (.996, .998).

The odds ratios associated with the individual predictors indicate that for one point (1 *SD*, 2 *SD*, and 3 *SD*) increase in UGPA, master's students were 1.56 (1.21, 1.45, and 3.53) times more likely to attain a degree within a 6-year time limit. One *SD* for UGPA was 0.42. For one point (1 *SD*, 2 *SD*, and 3 *SD*) increase in GREQ, master's students were 1.001(1.15, 1.32, and 1.51) times more likely to attain a degrees within a 6-year time limit. One *SD* for GREQ was 138.11. However, the odds ratio associated with GREV was significant but less than 1.00, *OR* = .997, 95% CI (996, 998), indicating that those with higher GREV scores were less likely to attain degrees within a 6-year time limit than those with lower GREV scores. This was not due to multicollinearity because the correlation of degree attainment with GREV was also significant and negative, r = ..10, *p* < .001. In addition, as shown in Table 14, the correlations of master's students' degree attainment within a 6-year time limit with GREV were all negative when evaluated by sex or race/ethnicity groups.

Results indicate that master's students' degree attainment within a 6-year time limit was associated with higher UGPA, higher GREQ, and lower GREV scores.

*Relationship between doctoral students' degree attainment status and the composite predictor.* The logistic regression equation of the composite predictor for doctoral students' degree attainment status within a 9-year time limit was significant,  $\chi^2$  (3, N = 400) = 12.69, p = .005, pseudo  $R^2$  = .030 to .044. As shown in Table 9,

between 3% and 4.4% of the variability in the doctoral students' degree attainment within

a 9-year time limit was explained by the composite predictor.

Table 9

Logistic Regression of Doctoral Students' Degree Attainment Status within a 9- Year Time Limit on the Composite Predictor for Research Question 3

Wald's								
Predictor	В	SE	χ2	df	р	OR	959	% CI
GREV	0.00017	0.001	.02	1	.879	1.000	0.998	1.002
GREQ**	0.00233	0.001	6.73	1	.009	1.002	1.001	1.004
UGPA*	0.65376	0.292	5.00	1	.025	1.923	1.084	3.409
Constant	-2.653	1.164	5.19	1	.023	0.070		

*Note.* GREV is GRE Verbal Reasoning Test scores. GREQ is GRE Quantitative Reasoning Test score. UGPA is undergraduate GPA. Model  $\chi^2$  (3, N = 422) = 12.69, p < .005. Cox and Snell  $R^2 = .030$ , Nagelkerke  $R^2 = .044$ .

p < .05. \*\* p < .01.

The coefficients of the three individual predictors showed that two predictors, GREQ and UGPA were significant (GREQ, p = .009, UGPA, p = .025) and positive (GREQ, B = .00233, UGPA, B = .65376) in predicting doctoral students' degree attainment within nine years. GREQ was a stronger predictor of degree attainment, OR =1.002, 95% CI (1.001, 1.004), than UGPA, OR = 1.92, 95% CI (1.08, 3.41). GREV was not a significant predictor, p = .879 in this model, OR = 1.00, 95% CI (.998, 1.002).

The odds ratios associated with the individual predictors indicate for one point (1 *SD*, 2 *SD*, and 3 *SD*) increase in in GREQ, doctoral students were 1.002 (1.34, 1.80, and 2.41) times more likely to attain degrees within a 9-year time limit. One *SD* for GREQ was 125.78. For one point (1 *SD*, 2 *SD*, and 3 *SD*) increase in UGPA, doctoral students were 1.92 (1.28, 1.65, and 2.12) times more likely to attain degrees within a 9-year time

limit. One *SD* for UGPA was 0.38. The odds ratio associated with doctoral students' GREV scores was not significant.

Results indicate that doctoral students' degree attainment within a 9-year time limit was associated with higher GREQ and higher UGPA scores.

In summary, for both master's and doctoral students, the logistic regression equations of degree attainment within the time limit on the composite predictor were significant (master's, p < .001, doctoral, p = .005). Between 1.5 to 2.5% (master's) and 3.0 to 4.5% (doctoral) of the variability in students' degree attainment within the time limit were explained by the linear combination of the composite predictor (GREV, GREQ, and UGPA). In addition, for both master's and doctoral students, degree attainment within the time limit was associated with higher UGPAs (master's OR = 1.56, doctoral OR = 1.92) and higher GREQ scores (master's OR = 1.001, doctoral OR =1.002).

# Research Question Four: Hierarchical Multiple Linear and Logistic Regression Analyses

The fourth research question was to evaluate whether the relationships between the composite predictor (GREV, GREQ, and UGPA) and success in graduate school, as measured by FGPA, CGPA, and degree attainment status differed by race/ethnicity and sex: more specifically, whether the relationships of the three success measures with the composite predictor differed between White students and minority students (e.g., African American, Asian American, and Hispanics), and between men and women. Note that international students were included in the analyses as one category of race/ethnicity group. For the master's degree level analysis, the Not Reported group was also included. Native Americans were not included in either master's or doctoral degree level analyses due to small sample sizes for all research variables. For doctoral degree level analysis, Asian Americans were also excluded due to a small sample size in the logistic regression analysis for degree attainment status variable.

# Relationships between the Three Success Measures and the Composite Predictor by Race/Ethnicity and Sex

To measure the differences by race/ethnicity and sex, hierarchical multiple linear regression analyses on FGPA and CGPA and hierarchical multiple logistic regression analyses on students' degree attainment status were performed. In all models, the composite predictor included GREV, GREQ, UGPA, sex, race/ethnicity, and interactions between each of the composite predictors and male, and interactions between each of the composite predictors and each race/ethnicity. For all models, The White student group was the reference group. The criterion variables were FGPA, CGPA, and degree attainment status.

To test for the significance of the interaction of race/ethnicity by the composite predictor, two regression models were fitted and tested for differences. In the first step, a multiple linear regression model (or a multiple logistic regression model) was constructed with main effects of the composite predictor, male, race/ethnicity, and cross products of male with each composite predictor. In the second step, cross products of race/ethnicity with each composite predictor were added. The  $R^2$  change (or the difference in the chi-square statistics) between the two models tested the significance of the interaction of race/ethnicity by the composite predictor.

Similarly, to test for the significance of the interaction of male by the composite predictor, in the first step, a multiple linear regression model (or a multiple logistic regression model) was constructed with main effects of the composite predictor, male, race/ethnicity, and cross products of race/ethnicity with each of the composite predictor, and in the second step, cross products of male with each of the composite predictors were added. The  $R^2$  change (or the difference in the chi-square statistics) between the two models tested the significance of the interaction of male by the composite predictor.

Differences in predictive relationships between FGPA and CGPA and the composite predictor among race/ethnicity and sex groups were examined in terms of significance (p - value) of the regression equation, magnitude of  $R^2$  change, and regression coefficients (B and  $\beta$ ) of the individual predictors. Differences in predictive relationships between degree attainment status and the composite predictor among race/ethnicity and sex groups were examined in terms of significance (p - value) of the logistic regression equation, pseudo  $R^2$  (Cox and Snell  $R^2$ , Nagelkerke  $R^2$ ), regression coefficients (B) of the individual predictors, odds ratios (ORs) and 95% confidence intervals for ORs.

# Relationships between FGPA and the composite predictor by race/ethnicity and sex.

Relationships between master's students' FGPA and the composite predictor by race/ethnicity and sex. The interaction of race/ethnicity by the composite predictor (GREV, GREQ, and UGPA) on master's students' FGPA was significant,  $R^2$  change = .004, F(15, 5,714) = 1.74, p = .038, indicating that the relationship between master students' FGPAs and the composite predictor were significantly different by race/ethnicity. Further investigation showed that the interaction of race/ethnicity by GREQ was significant,  $R^2$  change = .002, F(5, 5,714) = 2.64, p = .022. But closer examination of the interactions of each race/ethnicity group by GREQ revealed that none were significant in the model.

Neither the interactions of race/ethnicity by UGPA,  $R^2$  change = .001, F(5, 5,724) = 1.76, p = .118, nor the interactions of race/ethnicity by GREV,  $R^2$  change = .001, F(5, 5,719) = .805, p = .546, were significant. This indicates that the relationship between FGPA and the composite predictor was not significantly different by race/ethnicity.

The interaction of male by the composite predictor was not significant, the  $R^2$  change = .000, F(3, 5,726) = .886, p = .447, indicating that the relationship between master students' FGPA and the composite predictor was not significantly different by sex.

Results indicate that the relationships between master's students' FGPA and the composite predictor were not significantly different among African Americans, Asian Americans, Hispanics, International students, NRs or White students, nor were the relationships between master's students' FGPA and the composite predictor significantly different between men or women.

Relationships between doctoral students' FGPA and the composite predictor by race/ethnicity and sex. The interaction of race/ethnicity by the composite predictor on doctoral students' FGPA was significant, the  $R^2$  change = .019, F(15, 1300) = 1.75, p = .037, indicating that the relationships between doctoral students' FGPA and the composite predictor were significantly different by race/ethnicity. As shown in Table 10, further investigation showed that the interaction of race/ethnicity by GREQ on FGPA was significant,  $R^2$  change = .010, F (5, 1300) = 2.86, p = .014. Closer examination of the interactions of each race/ethnicity group by GREQ revealed that for NRs (those whose race/ethnicity was not reported, n = 24 out of 1377, or 1.7%), the relationship between FGPA and GREQ was significantly different from that of the White students. For NRs, higher FGPAs were significantly associated with higher GREQ scores,  $\beta$  = .58, p = .020, however, for White students, the relationship between FGPA and GREQ scores was not significant,  $\beta$  = -.011, p = .856. For the other groups in the doctoral degree level, African American, Asian American, Hispanic, and International students, the relationships between doctoral students' FGPA and GREQ were not significantly different from that of the White students after controlling for the composite predictor, male, and race/ethnicity.

Neither the interaction of race/ethnicity by UGPA,  $R^2$  change = .001, F (5, 1310) = .414, p = .839, nor the interaction of race/ethnicity by GREV,  $R^2$  change = .007, F (5, 1305) = 1.94, p = .083, was significant. This indicates that the relationship between FGPA and UGPA, and the relationship between FGPA and GREV were not significantly different by race/ethnicity. The interaction of male by the composite predictor was not significant, the  $R^2$  change = .003, F (3, 1312) = 1.48, p = .217, indicating that the relationships between doctoral students' FGPA and the composite predictor were not significantly different for men and women.

# Table 10

	First Year Graduate GPA						
Predictor	В	SE	β	p			
Constant	2.532	.255					
GREV**	.001	.000	.156	.000			
GREQ	00004	.000	011	.007			
UGPA***	.240	.057	.210	.856			
male **	067	.025	077	.000			
African Am.	.118	.433	.081	.007			
Asian Am.	.315	.753	.133	676			
Hispanic	.264	.354	.259	.456			
International	.443	.398	.465	.265			
NR	077	1.201	023	.949			
NR by GREQ**	.003	.0009	.547	.002			

*Hierarchical Multiple Regression Analysis of Doctoral Students' FGPA on the Composite Predictor for Research Question 4* 

*Note*. N = 1,325. Am. is an abbreviation of American. International is an abbreviation of international students. NR is a group whose race/ethnicity was not reported.

Only statistically significant coefficients are included in the table. The White students are the reference group. Model  $R^2 = .072$ , F(24, 1300) = 4.200, p < .001. \*\*p < .01. \*\*\*p < .001.

In summary, for both master's and doctoral students, the relationships between FGPA and the composite predictor were not significantly different for most of race/ethnicity groups after controlling for the composite predictor and sex. For doctoral students only, the relationship between FGPA and GREQ for NRs was significant and positive and significantly different from that of the White students after controlling for the composite predictor and sex. For both master's and doctoral students, the relationships between FGPA and the composite predictor (GREV, GREQ, and UGPA) for men and women were not significantly different from each other after controlling for the composite predictor and race/ethnicity. In addition, for both master's and doctoral students, lower FGPAs were associated with being male, indicating men's FGPAs were lower than those for women.

Relationships between CGPA and the composite predictor by race/ethnicity and sex.

*Relationships between master's students' CGPA and the composite predictor by race/ethnicity and sex.* The interaction of race/ethnicity by the composite predictor (GREV, GREQ, and UGPA) was not significant,  $R^2$  change = .001, F(15, 4630) = .47, p = .957. The interaction of male by the composite predictor was not significant,  $R^2$ change = .000, F(3, 4,627) = .41, p = .747. This indicates that the relationships between master's students' CGPA and the composite predictor were not significantly different either by race/ethnicity or sex.

Results indicate that neither the relationships between master's students' CGPA and the composite predictor were significantly different among African Americans, Asian Americans, Hispanics, NRs, International students or White students, nor were the relationships between master's students' CGPA and the composite predictor significantly different between men or women.

*Relationships between doctoral students' CGPA and the composite predictor by race/ethnicity and sex.* The relationships between doctoral students' CGPA and the composite predictor by race/ethnicity and sex were similar to those of the master's students. The interaction of race/ethnicity by the composite predictor was not significant,  $R^2$  change = .013, F(12, 972) = 1.15, p = .313. The interaction of male by the composite predictor was not significant,  $R^2$  change = .001, F(3, 969) = 0.26, p = .855. This indicates that the relationships between doctoral students' CGPA and the composite predictor were not significantly different either by race/ethnicity or sex.

Results indicate that neither the relationships between doctoral students' CGPA and the composite predictor were significantly different among African Americans, Asian Americans, Hispanics, International students, or White students, nor were the relationships between doctoral students' CGPA and the composite predictor significantly different between men or women.

In summary, for both master's and doctoral degree levels, the relationships between CGPA and the composite predictor were not significantly different either by race/ethnicity or by sex. Neither the relationships between students' CGPAs and the composite predictor were significantly different among African Americans, Asian Americans, Hispanics, International students, or White students, nor were the relationships between students' CGPAs and the composite predictor significantly different between men or women. In addition, for both master's and doctoral programs, lower CGPAs were associated with being male, indicating men's CGPAs were lower than those for women.

**Relationships between degree attainment status and the composite predictor by race/ethnicity and sex.** Hierarchical multiple logistic regression analyses were performed to measure differences in relationships between degree attainment status and the composite predictor (GREV, GREQ, and UGPA) by race/ethnicity and sex. The differences in relationships were examined in terms of significance (*p* - value) of the

regression model, pseudo  $R^2$  (Cox and Snell  $R^2$ , Nagelkerke  $R^2$ ), regression coefficients (*B*), odds ratios (*OR*) and 95% confidence intervals for the *OR*s.

Relationships between master's students' degree attainment status and the composite predictor by race/ethnicity and sex. The interaction of race/ethnicity by the composite predictor on master's students' degree attainment was significant,  $\chi^2 (15, N = 5,012) = 35.03, p = .002$ , indicating that the relationships between master's students degree attainment within a 6-year time limit and the composite predictor were significantly different by race/ethnicity. As shown in Table 11, further investigation showed that the interaction of race/ethnicity by UGPA, *Wald* = 13.48, *df* = 5, *p* = .019, and interaction of race/ethnicity by GREQ, *Wald* = 11.52, *df* = 5, *p* = .042, were significant. The interaction of race/ethnicity by GREV was not significant, *Wald* = 7.84, *df* = 5, *p* = .166.

Closer examination of the interaction of each race/ethnicity group by UGPA revealed that the interaction for International students was significant, B = 1.18, p < .001, OR = 3.25, 95% CI (1.72, 6.13). This indicates that for one point increase (1 SD, 2 SD, 3 SD) in UGPA, International master's students were 3.3 (1.52, 2.32, 3.53) times more likely to attain a degree within a 6-year time limit than White students (1SD increase for master's students' UGPA was 0.42). For other racial/ethnicity groups, the interactions of race/ethnicity by UGPA were not significant. That is, for African Americans, Asian Americans, Hispanics and NRs, the relationships between master's students' degree attainment within a 6-year time limit and UGPA were not significantly different from the White students.

Examination of the interaction of each race/ethnicity group by GREQ revealed that the interaction for Asian Americans was significant, B = -.004, p = .025, OR = .996, 95% CI (.993, .999). This indicates that for 1 point (1*SD*, 2 *SD*, 3*SD*) increase in GREQ,

White master's students were 1.004 (1.72, 3.03, 5.26) times more likely to attain a degree

within a 6-year time limit than Asian American students (1SD was 138.11).

Table 11

*Hierarchical Multiple Logistic Regression Analysis of Master's Students' Degree Attainment Status within a 6-Year Time Limit on the Composite Predictor for Research Question 4* 

			Wald's				95% (	CI for OR
Predictor	В	SE	χ2	df	р	OR	LL	UL
GREV***	004	.001	20.15	1	< .001	.996	.995	.998
GREQ	.001	.001	.09	1	.759	1.000	.999	1.002
UGPA	.107	.177	.37	1	.545	1.113	.787	1.576
Male	-1.181	.712	2.75	1	.097	.307	.076	1.239
Ethnicity**			16.76	5	.005			
African American	-1.827	1.142	2.56	1	.110	.161	.017	1.510
Asian American	.902	1.876	.23	1	.630	2.465	.062	97.375
Hispanic	-1.541	.882	3.05	1	.081	.214	.038	1.206
International***	-4.813	1.283	14.07	1	< .001	.008	.001	1.00
NR	-4.308	3.677	1.37	1	.241	.013	.000	18.142
Ethnicity by UGPA*			13.48	5	.019			
Int'l by UGPA***	1.178	.324	13.23	1	< .001	3.249	1.722	6.131
Ethnicity by GREV			7.84	5	.166			
Ethnicity by GREQ*			11.52	5	.042			
Asian Am. by GREQ*	004	.002	5.03	1	.025	.996	.993	.999
Constant	2.976	.731	16.59	1	< .001	19.616		

*Note.* N = 5,015. Int'l (international) is an abbreviation of international students. NR is a group whose race/ethnicity was not reported. Ethnicity was used for race/ethnicity for brevity. Only statistically significant interaction coefficients were included in the table. The White students are the reference group. Model  $\chi^2$  (27, N = 5012) = 137.12, p < .001. Cox and Snell  $R^2 = .027$ , Nagelkerke  $R^2 = .046$ . \* p < .05. \*\* p < .01. \*\*\* p < .001.

That is, Asian American students were less likely to attain degrees within a 6-year time limit than White students.

The interaction of male by the composite predictor was not significant,

 $\chi^2$  (3, N = 5,012) = 6.75, p = .080. This indicates that the relationships between master's students' degree attainment within a 6-year time limit and the composite predictor were not significantly different for men and women.

Results indicate that the relationships between master's students' degree attainment within a 6-year time limit and the composite predictor by race/ethnicity were significantly different for two individual predictors (UGPA, and GREQ): International students with higher UGPAs were more likely to attain degrees within a 6-year time limit than White students. In addition, White students with higher GREQ scores were more likely to attain degrees within a 6-year time limit than Asian American students. The relationship between master's students' degree attainment within a 6-year time limit and the composite predictor was not significantly different between men and women students.

Relationships between doctoral students' degree attainment status and the composite predictor by race/ethnicity and sex. The interaction of race/ethnicity by the composite predictor was not significant,  $\chi^2 (9, N = 400) = 10.28, p = .328$ . The interaction of male by the composite predictor was not significant,  $\chi^2 (3, N = 400) = 0.61$ , p = .894. This indicates that the relationship between doctoral students' degree attainment within a 9-year time limit and the composite predictor was not significantly different either by race/ethnicity or sex.

In summary, for master's students, the relationships between degree attainment within a 6-year time limit and two individual predictors (UGPA and GREQ) by race/ethnicity were significantly different. International students with higher UGPAs were more likely to attain degrees within a 6-year time limit than White students; White students with higher GREQ scores were more likely to attain degrees within a 6-year time limit than Asian American students.

For doctoral students, the relationships between degree attainment status within a 9-year time limit and the composite predictor by race/ethnicity were not significantly different among African American, Hispanics, International students and the White students. For both master's and doctoral students, the relationships between degree attainment status within a time limit and the composite predictor were not significantly different between men and women students.

#### **Summary**

The first research question evaluated the predictive relationship between GREV, GREQ, and GRET and success in graduate school, as measured by three success measures, FGPA, CGPA, and degree attainment status. Results showed that for master's students, higher FGPA scores were associated with higher GREV (r = .15), higher GREQ scores (r = .05), and higher GRET scores (r = .12). Higher CGPAs were associated with higher GREV (r = .09). Master's students' degree attainment within a 6-year time limit was negatively associated with GREV (r = .10) and GRET scores (r = .05).

For doctoral students, higher FGPAs were associated with higher GREV (r = .07) and higher GRET scores (r = .07). Higher CGPAs were associated with higher GRET scores (r = .07). Doctoral students' degree attainment within a 9-year time limit was associated with higher GREQ scores (r = .14) and higher GRET scores(r = .13).

The second research question evaluated the predictive relationship between undergraduate GPA (UGPA) and success in graduate school, as measured by the three success measures, FGPA, CGPA, and students' degree attainment status. Results showed that for master's students, higher FGPAs were associated with higher UGPAs (r = .21). Higher CGPAs were associated with higher UGPAs (r = .17). Master's students' degree attainment within a 6-year time limit was associated with higher UGPAs (r = .06). For

doctoral students, the relationships between FGPA, CGPA, degree attainment within a time limit and UGPA were similar to those found for master's students. Higher FGPAs were associated with higher UGPAs (r = .19). Higher CGPAs were associated with higher UGPAs (r = .13). Doctoral students' degree attainment within a 9-year time limit was associated with higher UGPAs (r = .12).

The third research question evaluated the predictive relationship between the composite predictor (GREV, GREQ, and UGPA) and success in graduate school, as measured by three success measures, FGPA, CGPA, and degree attainment status. Results showed that all regression equations of the three success measures on the composite predictors were significant for both master's and doctoral degree levels.

The regression equations of FGPA on the three composite predictors were significant for the two degree levels with 4 % (doctoral) and 6% (master's) of the variability in students' FGPAs predicted by the linear combination of the composite predictor, ps < .001. The regression equations of CGPA on the composite predictors were significant for the two degree levels with 2 % (doctoral) and 4% (master's) of the variability in students' CGPAs predicted by the linear combination of the composite predictor, ps < .001. The logistic regression equations of degree attainment within the time limit on the composite predictors were significant for the two degree predictors were significant for the two degree levels with 3% to 4.4 % (doctoral) and 1.5 % to 2.5 % (master's) of the variability in students' degree attainment within a time limit were predicted by the linear combination of the composite predictor, p < .001 (master's) and p < .005 (doctoral).

Examination of the individual predictors revealed that UGPA was the predictor that consistently contributed to predicting all three success measures, FGPA (master's  $\beta$  =

.20, doctoral  $\beta$  = .19), CGPA (master's  $\beta$  = .16, doctoral  $\beta$  = .13), and degree attainment within a time limit (master's OR = 1.56, doctoral OR = 1.92). It is noteworthy that for the degree attainment status variable, GREQ was also significant (master's OR = 1.001, doctoral OR = 1.002) in addition to UGPA in predicting students' degree attainment within a time limit.

The fourth research question evaluated whether the relationships between the composite predictor (GREV, GREQ, and UGPA) and success in graduate school, as measured by FGPA, CGPA, and degree attainment status, differed among White students, African American, Asian American, Hispanics, NRs, and International students, and between men and women. For master's students, the relationships between FGPA and the composite predictor were not found to be significantly different among any of the groups.

For doctoral students, the relationship between FGPA and GREQ was significantly different between NRs and White students. For NRs, higher FGPAs were significantly associated with higher GREQ scores,  $\beta = .58$ , p = .020. However, for White students, the relationship between FGPA and GREQ scores was not significant,  $\beta = -.011$ , p = .856. No other relationships were significant.

For both master's and doctoral students, the relationships between CGPA and the composite predictor were not significantly different either by race/ethnicity or by sex.

For master's students, the relationship between degree attainment and UGPA was significantly different between International students and White students. For one point increase in UGPA, International master's students were 3.3 times more likely to attain a degree within a 6-year time limit than White students, OR = 3.25, 95% CI (1.72, 6.13).

The relationship between master's students' degree attainment and GREQ was significantly different between Asian Americans and White students. For one standard deviation (1*SD* was 138.11) increase in GREQ, White students was 1.72 times more likely to attain a degree within a 6-year time limit than Asian Americans. The relationship between master's students' degree attainment within a 6-year time limit and the composite predictor was not significantly different between men and women students.

For doctoral students, the relationships between degree attainment within a 9-year time limit and the composite predictor were not significantly different either by race/ethnicity or by sex.

Chapter V discusses the findings of the research study, implications for theory and practice, followed by limitations of the study.

#### **CHAPTER V**

#### DISCUSSION

Chapter 5 presents discussion of the findings, implications for theory, practice, research, and limitations of the study. The chapter ends with conclusions.

## **Summary of the Study**

Most graduate schools require applicants to submit GRE General Test scores (GRE Verbal Reasoning Section Test scores [GREV] and GRE Quantitative Reasoning Section Test scores [GREQ]) for admission considerations. The rationale for the score requirement is based on utility that the test publisher proposed, i.e., the "GRE General Test scores are valid predictors of success in the first year of graduate school for all students" (ETS, 2009, p. 7). Graduate schools use the test score to evaluate whether an applicant is ready for graduate level study, and to predict further whether the applicant is capable of progressing through a program, and ultimately attaining a degree.

However, admission tests, including GRE, have long been viewed as a major barrier to higher education for minority students (Zwick, 2002). An implication from this view is that the rationale for the use of a test is not applicable to minority students. That is, for a graduate admission test, the GRE General Test scores are not valid predictors of first year graduate GPA (FGPA) for minority students. In other words, the relationships between the GRE General Test scores and FGPA differ between White (majority of the population) students and minority students. A problem is there is not sufficient empirical evidence to support this view: GRE validation studies based on minority students have been scarce mainly due to small sample sizes; and the findings have been inconsistent. The purpose of the present study was to examine the predictive validity of the GRE scores and to determine whether the predictive relationships differ by race/ethnicity and sex. The predictors included GRE Verbal Reasoning Test scores, GRE Quantitative Reasoning Test scores, and Undergraduate GPA (UGPA). The criterion was success in graduate school, as measured by first year graduate GPA (FGPA), cumulative graduate GPA (CGPA), and degree attainment status.

The research questions that guided the study were as follows:

- To what degree do GRE scores taken singly (individual GREV and GREQ) and in total, sum of GREV and GREQ (GRET), predict success in graduate school, as measured by FGPA, CGPA, and degree attainment status?
  - a. To what degree does GREV predict success in graduate school?
  - b. To what degree does GREQ predict success in graduate school?
  - c. To what degree does GRET predict success in graduate school?
- 2. To what degree does UGPA predict success in graduate school, as measured by FGPA, CGPA, and degree attainment status?
- 3. To what degree do GREV, GREQ, and UGPA, as a composite predictor, predict success in graduate school, as measured by FGPA, CGPA, and degree attainment status?

4. Do the relationships between GREV, GREQ and UGPA, as a composite predictor, and success in graduate school, as measured by FGPA, CGPA, and degree attainment status, differ by race/ethnicity and sex?

Both master's and doctoral student records of 10 academic years, from 2000 to 2010, were used to investigate the relationships among the GREV, GREQ, UGPA, and

success in graduate school, as measured by FGPA, CGPA, and degree attainment status. Pearson's correlation, multiple linear (and logistic) regression, hierarchical multiple linear (and logistic) regression analyses were performed to answer the four research questions.

Results of statistical analyses differ by degree level. Correlational analyses indicated that the GREV and GREQ scores, both individually, and in total, were valid predictors of first year graduate GPA for master's students. For doctoral students, individual GREV and the GRE total scores, but not individual GREQ, were valid predictors of first year graduate GPA. Regression and correlational analyses indicated that UGPA was a consistent predictor of all three success measures, FGPA, CGPA, and degree attainment for both master's and doctoral degree levels.

Results of multiple linear and logistic regression analyses indicated that the composite predictor (GREV, GREQ, UGPA) predicted FGPA, CGPA, and degree attainment for both degree levels. The results also indicated that the relationships of the composite predictor (GREV, GREQ, and UGPA) with the three success measures were not significantly different among African American, Hispanic, and White students for either masters or doctoral degree level. However, the relationship between master's degree attainment and GREQ for Asian Americans, the relationship between master's degree attainment and UGPA for International students, and the relationships between doctoral FGPA and the GREQ for NRs were significantly different from that of White students.

# **Discussion of the Results**

This section discusses the results of each research questions. This was a test validation study whose main attention was concentrated on checking the accuracy of the specific prediction proposed by the test publisher across a large racially and ethnically diverse student population (Cronbach, 1971), i.e., to check whether the GRE scores are valid predictors of first year graduate GPA (FGPA) for all students. The predictive relationships between GRE scores and FGPA (and other success measures) were examined through the research questions.

#### **Research Question One**

The first research question evaluated the relationships between GRE scores (GREV, GREQ, and GRET) and success in graduate school, as measured by FGPA, CGPA, and degree attainment status.

# **Relationship between GRE Scores and FGPA**

Results indicated that the relations between GRE scores (GREV, GREQ, and GRET) and FGPA differed by degree level. For master's level students, findings were consistent with the proposed prediction by the ETS that the GRE scores are valid predictors for FGPA. Although the magnitudes of the correlations were small, master's students who had higher FGPA tended to have higher GREV (r = .15), higher GREQ (r = .05), and higher GRET scores (r = .12). While findings of the present study are not directly comparable to those of the previous studies because most GRE validation studies were conducted based on total graduate student population, not differentiating master's from doctoral degree level, they are consistent with most of meta-analytic, and large

studies (Goldberg & Alliger, 1992; Kuncel et al., 2001; Schneider & Briel, 1990; Thacker & Williams, 1974; Wilson, 1979).

For doctoral level students, the findings partially supported the proposition of ETS that GRE scores are valid predictors for FGPA. Although FGPA was significantly correlated with GREV (r = .07) and GRET (r = .07), it was not significantly correlated with GREQ. Nevertheless, this indicates that doctoral students who had higher FGPAs tended to have higher GREV and higher GRET scores.

Overall, meta-analytic and large studies found significant relationships between the GRE scores and FGPA. The correlations between FGPA and GREV ranged from .22 to .29, the correlation between FGPA and GREQ ranged from .15 to .28, the correlation between FGPA and GRET ranged from .18 to .34. While small correlations typical in GRE validation studies were expected due to a ceiling effect, e.g., restriction of range of scores (Ingram, 1983; Kuncel et al., 2001; Mehrens & Lehmann, 1987; Schmidt & Hunter, 1981; Schneider & Briel, 1990), criterion unreliability((Kuncel et al., 2001; Mehrens & Lehmann, 1987; Schmidt & Hunter, 1981), and compensatory selection(Ingram, 1983; Schneider & Briel, 1990; Whitworth & Barrientos, 1990), the present study found smaller correlations than those in the previous studies.

The reasons for smaller correlations found in the present study may include the following. First, more restriction of range of scores may have reduced the correlation for this sample. That is, variances among students' GRE scores and FGPAs in this sample were smaller than those of the population reported by the ETS (2009 - 2010): the *SD* for GREV was 121, and the *SD* for GREQ was 152. For this sample, the *SD*s for GREV were 100.98 (master's), 107.50 (doctoral), respectively, and the *SD*s for GREQ were 138.11

(master's), 125.78 (doctoral), respectively. The standard deviations for both master's and doctoral degree levels from the sample were smaller of those of the total GRE test examinees, which may have influenced the correlations.

Second, compensatory selection may have reduced the correlation between FGPA and the GRE scores for this sample. That is, if a student with low GRE scores was offered an admission because of his (her) high UGPA, then, their FGPA and GRE scores are not related. Thus, compensatory selection of this type reduces the correlation between FGPA and the GRE scores.

#### **Relationship between GRE Scores and CGPA**

As discussed in Empirical Review in Chapter 2, ETS did not propose that the GRE scores would predict CGPA. However, CGPA, is included as a criterion measure of the present study because it is used as an interim measure of success in graduate school and one of the most frequently used variables in GRE validation studies.

Results indicated that the relationships between GRE scores (GREV, GREQ, and GRET) and CGPA differed by degree level. For master's students, the relationships between CGPA and GRE scores (GREV, GREQ, and GRET) were similar to those between FGPA and GRE scores, but with lower correlations (r = .10, .06, .and 09, respectively). Master's students who had higher CGPA tended to have higher GREV, higher GREQ, and higher GRET scores. This finding is consistent with the largest meta-analytic study (Kuncel et al., 2001) in which the researchers found the relationships between CGPA and GRE scores were similar, but to a lesser degree, to those between FGPA and GRE scores.

For doctoral students, results indicated that doctoral students who had higher CGPA tended to have higher GRET scores (r = .07). Findings of the present study are consistent with those of Burton and Wang (2005), while they contradict with those of Thacker and Williams (1974). Burton and Wang found larger correlation between graduate GPA (GGPA) and GRET than those of individual GREV, and GREQ. The correlation between GGPA and GRET ranged from .29 to .39 while the correlation between GGPA and GREV ranged from .06 to .20. The correlation between GGPA and GREV ranged from .06 to .20. The correlation between GGPA and GREV ranged from .06 to .20. The correlation between GGPA and GREV ranged from .06 to .20. The correlation between GGPA and GREQ ranged from .01 to .19. However, Thacker and Williams found doctoral students' CGPAs were significantly and positively associated with individual GREV and GREQ scores, but not with GRET (n = 252 doctoral students). Further study is needed to determine the utility of the GRET scores in predicting CGPA.

#### **Relationship between GRE Scores and Degree Attainment Status**

Results indicated that the relationships between GRE scores and degree attainment status differed by degree level. For master's students, those who attained a degree within a 6-year time limit tended to have lower GREV, GREQ and lower GRET scores as compared to those who did not attain a degree. This result may be a reflection of (a) suppressor variable effect or (b) a variable not measured in the study (Cohen et al., 2003). This finding contradicts previous meta-analytic study (Kuncel et al., 2001) in which the researchers found graduate students who had higher GREV (*r* ranged from .14 to 34) and higher GREQ scores (*r* ranged from .08 to 26) tended to attain a degree within a given time limit as compared to those who had not. This interesting result warrants further research. For doctoral students, those who had higher GREQ and higher GRET scores tended to attain a degree within a 9-year time limit as compared to those who had not. This finding is consistent with a study (House, 1997) based on American Indian students (n = 28: masters, 26; doctoral students, 2). The researcher found that degree completion was significantly and positively associated with only GREQ scores (r = .47). Note that correlation of this magnitude is very rare in GRE validation studies, and the finding was based on a very small sample (n = 28). Findings of the present study and House's suggest that further studies would be useful to investigate the characteristics of GREQ in relation to prediction of degree attainment.

# **Research Question Two**

The second research question evaluated the relationships between undergraduate GPA (UGPA) and success in graduate school, as measured by FGPA, CGPA, and students' degree attainment status.

#### **Relationship between UGPA, and FGPA, CGPA, Degree Attainment Status**

Results indicated that the relationships between all three success measures (FGPA, CGPA, and degree attainment) and UGPA were significant for both master's and doctoral degree levels. This indicates that graduate students who had higher FGPAs, higher CGPAs, and attain a degree within a given time limit tended to have higher UGPAs. For both master's and doctoral degree levels, the largest correlation was between FGPA and UGPA (master's, r = .21; doctoral, r = .19). The correlation between the three success measures and UGPA decreased in magnitude as the time passed farther away from the first year of graduate school. This result makes sense because as a student

progresses through a graduate program, fewer skills and less knowledge from undergraduate program may be related to the students' academic performance.

Findings of the present study are consistent with those of the largest meta-analytic study (Kuncel et al., 2001; n = 82,659) in which the researchers found all three success measures were associated with UGPA with decreasing size of correlations. The largest correlation was between FGPA and UGPA (r = .30), followed in order by the correlation between GGPA and UGPA (r = .28), and the correlation between degree attainment status and UGPA (r = .12).

## **Research Question Three**

The third research question evaluated the predictive relationship between the composite predictor (GREV, GREQ, and UGPA) and success in graduate school, as measured by three success measures (FGPA, CGPA, and degree attainment status).

# **Relationship between the Composite Predictor and FGPA**

Results of multiple linear regression analyses were similar for both the master's and doctoral degree level. The relationships between FGPA and the composite predictor were significant for both degree levels. For the master's degree level, about 6% of the variability in the master's students' FGPA was predicted by the linear combination of the composite predictor. UGPA ( $\beta$  = .20) and GREV ( $\beta$  = .13) were the individual variables that significantly contributed to predicting master's students' FGPA. Master's students who had higher FGPAs were more likely to have higher UGPAs and higher GREQ scores. This finding is inconsistent with the results of correlational analyses of the present study which indicated that taken individually, GREV, GREQ, and UGPA were significantly associated with FGPA.

For doctoral degree level, about 4% of the variability in the doctoral students' FGPA was predicted by the linear combination of the composite predictor. However, UGPA ( $\beta$  = .13) was the only variable that significantly contributed to predicting doctoral students' FGPA. This indicates that doctoral students who had higher FGPAs were more likely to have higher UGPAs. This finding is inconsistent with the results of correlational analyses of the present study which indicated that taken individually, GREV, GRET, and UGPA were significantly associated with FGPA. There are no comparable studies to evaluate the findings of the present study.

# **Relationship between the Composite Predictor and CGPA**

Results of the multiple linear regression analyses were similar for both the master's and doctoral degree level. The relationships between CGPA and the composite predictor were significant for both degree levels. For master's degree level, about 4% of the variability in the master's students' CGPA was predicted by the linear combination of the composite predictor. All three variables significantly contributed to predicting master's students' CGPA, indicating that master's students who had higher CGPAs were more likely to have higher GREVs ( $\beta = .07$ ), higher GREQs ( $\beta = .04$ ), and higher UGPAs ( $\beta = .16$ ). This finding is consistent with the correlational analyses of the present study where CGPAs were significantly and positively associated with GREV, GREQ, and UGPA. There are no comparable studies to evaluate the findings of the present study.

For doctoral degree level, about 2% of the variability in CGPA was predicted by the linear combination of the composite predictor. However, UGPA was the only variable that significantly contributed to predicting doctoral students' CGPA. This indicates that Doctoral students who had higher CGPAs were more likely to have higher UGPAs

 $(\beta = .13)$ . This finding is consistent with that found by Whitworth and Barrientos (1990). The researchers found that UGPA was the only predictor that contributed to predicting graduate students' CGPA (n = 952). GRE scores were not significant predictors in that study.

# **Relationship between the Composite Predictor and Degree Attainment Status**

Results of the multiple logistic regression analyses were significant for both master's and doctoral degree level. For master's degree level, between 2% and 3% of the variability in degree attainment were predicted by the composite predictor. Two variables, UGPA and GREQ, significantly contributed to predicting master's students' degree attainment within a 6-year time limit. This indicates that master's students who attained a degree within a given time limit were more likely to have higher UGPA (OR = 1.6) and higher GREQ scores (OR = 1.001). For doctoral degree level, between 3% and 4% of the variability in degree attainment were predicted by the linear combination of the composite predictor. Similar to the master's degree level, the two variables, UGPA (OR = 1.9) and GREQ (OR = 1.002), significantly contributed to predicting doctoral students' degree attainment within a 9-year time limit. There are no directly comparable studies to evaluate this finding. However, it would be useful that future researchers investigate the characteristics of UGPA and GREQ in relation to degree attainment status.

#### **Research Question Four**

The fourth research question evaluated whether the relationships of the three success measures to the composite predictor differ among White students, minority students (e.g., African American, Asian American, and Hispanics), and International students, or between men and women.
# Relationship between FGPA and the Composite Predictor (GREV, GREQ, UGPA) by Race/Ethnicity and Sex

Results of the hierarchical multiple linear regression analyses differed by degree level. For master's students, the relationships between FGPA and the composite predictor were not significantly different for African Americans, Asian Americans, Hispanics, International students, NRs or White students.

While several previous research studies compared the differences in relationships between GRE scores and graduate GPA among racial/ethnic groups using analysis of variance (ANOVA), there are no known studies that examined the differences in relationships between GRE scores and FGPA by racial/ethnic group using hierarchical multiple linear regression analysis as used for the present study. Hence, there are no comparable studies to evaluate this finding. In fact, most of the findings in Research Question Four are new to a GRE validation study.

More research studies are needed to investigate the relationships between GRE scores and FGPA by race/ethnicity group to resolve the issue as to whether the relationships differ among race/ethnicity groups. Existing literature has not provided a definitive answer to this question as of yet. Using hierarchical multiple linear regression analysis would be useful because unlike ANOVA, the hierarchical multiple linear regression analysis tests the relationship (slope) with the interactions of the predictor variable by race/ethnicity group. One should note that there are several differential prediction studies that used the hierarchical multiple linear regression analysis method involving SAT and college GPA (Young, 2001).

With respect to a sex difference, the relationships between master's students' FGPA and the composite predictor were not significantly different between men and women. This finding is consistent with those found by Sternberg and Williams (1997) in which the researchers found no significant differences in relationships of GRE scores with FGPA, or second year graduate GPA between men and women.

For doctoral students, the relationship between FGPA and GREQ score was significantly different between NRs (a group of students whose race/ethnicity was not reported, n = 24 out of 1377, or 1.7%) and the White students. NR students who had higher FGPAs were more likely to have higher GREQ scores than White students. Because no comparable studies are available to evaluate this finding, and the sample size of this group is small, this finding may only be valid for the particular group of students included in the study.

With respect to a sex difference, the relationships between doctoral students' FGPA and the composite predictor were not significantly different between men and women. The findings relating to sex differences examined through Research Question 4 were similar for all the relationships. There were no significant differences in the relationships between the composite predictor and the three success measures or between men and women for either master's or doctoral degree level.

# Relationship between CGPA and the Composite Predictor (GREV, GREQ, UGPA) by Race/Ethnicity and Sex

Results of the hierarchical multiple linear regression analyses were similar for both master's and doctoral degree levels. The relationships between CGPA and the composite predictor were not significantly different among African Americans,

Hispanics, International students, and White students for either degree level. Nor were the relationships between CGPA and the composite predictor significantly different between men and women. Further research is needed to validate these findings.

# Relationship between Degree Attainment Status and the Composite Predictor (GREV, GREQ, UGPA) by Race/Ethnicity and Sex

Results of the hierarchical multiple logistic regression analyses differed by degree levels. For master's students, the relationship between degree attainment status and the composite predictor by race/ethnicity was significantly different for two individual predictors (UGPA and GREQ): International students with higher UGPAs were more likely to attain a degree within a 6-year time limit than White students; and White students with higher GREQ scores were more likely to attain a degree within a 6-year time limit than Asian American students. For doctoral students, the relationships between degree attainment status and the composite predictor were not significantly different among African Americans, Hispanics, International students, and White students. The relationship between degree attainment status and the composite predictor by sex was not significantly different between men and women for either degree level. Further research is needed to validate the findings.

#### **Implications for Theory, Research, and Practice**

The validity of the GRE scores and UGPA for predicting success in graduate school has been of utmost interest to test takers, graduate school administrators, researchers, and test publishers because it involves a high-stakes decision. This section addresses implications for theory, research, and practice.

#### **Implications for Theory**

Results of the present study provide empirical evidence to support two major theories. They include performance and knowledge acquisition theory. According to performance theory (Campbell et al., 1993), an individual's various performances are viewed as a function of only three major determinants – declarative knowledge (also known as verbal information), procedural knowledge and skill (also known as intellectual skills), and motivation. The GRE General Test measures two of these major determinants, verbal information and intellectual skills, by Verbal Reasoning, and Quantitative Reasoning tests that are considered prerequisites for graduate study. Although it is not known exactly what percentage of variability in graduate students' academic performance should be accounted for by the two of the three determinants, significant correlations between the two determinants and FGPA demonstrate the validity of the principles of performance theory.

Second, findings of the present study also provided evidence to support a theory of knowledge acquisition. According to the knowledge acquisition theory (Gagne, 1962; Gagne & Medsker, 1996), learning is cumulative and hierarchical. That is, acquisition of knowledge occurs hierarchically from simpler level skills to more complex higher level skills by incorporating simpler level skills into increasingly more complex level skills. Theoretically, those who have acquired more prerequisite skills perform better than those who have not. The correlations found between FGPA and UGPA (master's, r = .21; doctoral, r = .19), and between CGPA and FGPA (master's, r = .49; doctoral, r = .58) in both degree levels demonstrate the validity of the principles of knowledge acquisition theory.

#### **Implications for Research**

Findings of the present study indicated a need for future research in several areas. First, it is recommended that future researchers analyze data by degree levels where applicable. Second, it is recommended that future researchers examine differential prediction among racial/ethnic groups using a hierarchical regression analysis. Third, it is recommended that future researchers examine the relationship between GRET and success in graduate school, as measured by, e.g., FGPA, CGPA, and degree attainment. Fourth, it is recommended that future researchers examine the variables, GREQ and UGPA in relations to degree attainment status. Lastly, it is recommended that future researchers examine the relationship between GRE scores and success in graduate school by discipline or department.

Examination of the predictive relationship of GRE Scores by degree level.

Most previous GRE validation studies were based on a total graduate student population. However, evidence from this study indicated that the relationships between GREV, GREQ, UGPA, and the three success measures differ by degree level. These differences by degree levels warrant further study.

**Studies of differential prediction.** Admission tests, including GRE, have long been viewed as a major barrier to higher education for minority students (Zwick, 2002). An implication from this view is that the rationale for the use of a test is not applicable to minority students. In other words, the relationships between the GRE General Test scores and FGPA differ between White (majority of the population) students and minority students. However, due to inconsistent findings and scarcity of empirical evidence, we cannot give definitive answer as to whether this is true.

Findings of the present study indicated that the relationship between master's degree attainment and GREQ for Asian Americans, the relationship between master's degree attainment and UGPA for International students, and the relationship between doctoral FGPA and GREV for NRs were significantly different from those for White students. However, this evidence is not sufficient to allow for any generalization. More studies of this type are needed to validate the findings and to better understand differential predictions.

**Examination of GRET score in relations to success in graduate school.** Correlational analyses yielded interesting results for this sample. All three success measures were significantly and positively associated with doctoral students' GRET scores as well as UGPA. While the magnitudes of the correlations of FGPA (r = .07) and CGPA (r = .07) to GRET were smaller than those of FGPA (r = .19) and CGPA (r = .13) to UGPA, the correlations of degree attainment to GRET (r = .13) and to UGPA (r = .12) were similar.

Most graduate schools use GRET scores for a preliminary screening measure without any valid empirical evidence to support the practice. Thus, further research on the validity of the variable, GRET would be needed.

#### Examination of the GREQ and UGPA in relation to degree attainment

**status.** Results of the regression analyses indicated that GREQ and UGPA significantly contributed to predicting degree attainment within a given time limit for both degree levels for the sample. This finding is unique to this sample. There are no known studies to compare to these findings. Future research will be useful to understand the relationships of GREQ and UGPA to degree attainment status.

#### Examination of the predictive relationship of GRE Scores by field or

**department**. Some of the previous research studies reviewed in the present study also investigated the predictive relationships of GRE scores to success in graduate school by field or department, and reported three types of information. First, a certain ethnic group tended to choose certain major fields (Burton & Wang, 2005; Sampson & Boyer, 2001). Second, GREV scores tended to be associated with the verbally-oriented fields, and the GREQ scores, with the quantitatively-oriented fields (Wilson, 1979). Third, higher correlations between FGPA and the GRE scores were found for departments whose students had higher GRE scores (Schneider & Briel, 1990). Based on these findings, future studies examining the predictive relationships of GRE scores to success in graduate school by field or department would yield some useful information that may be helpful in determining strengths of applicants in their chosen fields or departments.

#### **Implications for Practice**

The study has four types of practical implications that may affect educational decision making processes that use UGPA and GRE scores. While evidence of the present study supports the use of both UGPA and GRE scores in admission considerations, it will be helpful if graduate school administrators are informed of the study's specific findings to use as a reference in making educational decisions. First, UGPA consistently predicted all three success measures, FGPA, CGPA, and degree attainment status, for both master's and doctoral degree levels. Second, GREQ and UGPA predicted a degree attainment status within a given time limit for both master's and doctoral degree levels. Third, differential predictions were found for Asian Americans and International students. Fourth, no differential predictions were found

among African American, Hispanic, and White students for either master's or doctoral degree level in the relationships involving the three success measures and the composite predictor. Fifth, no differential predictions were found between men and women for either master's or doctoral degree level.

UGPA, the predictor for FGPA, CGPA, and degree attainment. Results of both correlational and regression analyses indicated that UGPA was positively and significantly associated with all three success measures for both masters and doctoral degree levels. This finding is consistent with most of large and meta-analytic studies (Burton & Wang, 2005; Kuncel et al, 2001; Schneider & Briel, 1990; Whitworth & Barrientos, 1990; Willingham, 1974; Wilson; 1979), which provides graduate school administrators with empirical and professionally defensible evidence to support the current practice of using undergraduate GPA for admission considerations.

**GREQ and UGPA, the predictors for degree attainment.** Results of regression analyses indicated that GREQ and UGPA significantly and positively contributed to predicting degree attainment for both master's and doctoral degree levels. Although there are no comparable studies to evaluate the findings, they are meaningful because they were based on the university's own student records of past 10 academic years (from 2000 to 2010) with a large sample (master's, n = 5,021, doctoral, n = 400). These statistically significant results will be useful for graduate school administrators as a reference in making educational decisions involving use of GRE scores and UGPA.

**No significant differences among African American, Hispanic, and White students.** Results of hierarchical linear and logistic regression analyses indicated that the relationships between FGPA, CGPA, and the composite predictor (GREV, GREQ, and

UGPA) were not statistically significantly different among African American, Hispanic, and White students for either master's or doctoral degree level. This indicates the degree of associations (slope weights for the predictors) between the predictors and the criterion variable) for each of these ethnic groups were not significantly different. This finding is contrary to a long-held view suggesting that the validity of the GRE scores for predicting minority students' success in graduate school may differ from that for White students (Dawes, 1971; ETS, 2008). This information will be useful for graduate school administrators as a reference in making admission decisions involving use of GRE scores and UGPA.

**No significant differences between men and women.** Results of hierarchical linear and logistic regression analyses indicated that the relationships between the three success measures (FGPA, CGPA, and degree attainment status) and the composite predictor (GREV, GREQ, and UGPA) were not statistically significantly different between men and women for either master's or doctoral degree level. This indicates the degree of associations (slope weights for the predictors) between the predictors and the criterion variable for men and women were not significantly different. This information will also be useful for graduate school administrators as a reference in making admission decisions involving use of GRE scores and UGPA.

**Differential prediction for Asian Americans, and International students.** Two pieces of empirical evidence for differential prediction were found at master's degree level which will be useful for reference in admission considerations. First, the relationship between degree attainment and GREQ differed significantly between Asian American and White students at the master's degree level. White students with higher

GREQ were more likely to attain a degree within a 6-year time limit than Asian American students. More specifically, for 1*SD* increase in GREQ (1SD = 138.11), White students were 1.7 times more likely to attain a degree within a 6-year time limit than Asian American students. This information can be useful in evaluating Asian American students in admission considerations. For the other groups in the master's degree level, African American, Hispanic, International student, and White student, the relationships between degree attainment and GREQ were not significantly different.

Second, the relationship between degree attainment and UGPA differed significantly between International students and White students at the master's degree level. International students with higher UGPA were more likely to attain a degree within a 6-year time limit than White students. More specifically, for 1*SD* increase in UGPA (1SD = 0.42), International students were 1.5 times more likely to attain a degree within a 6-year time limit than White students. This information can also be useful for reference in evaluating International students in admission considerations. All in all, the findings of the present study provide graduate school administrators with institution-specific validity data for UGPA and the GRE scores, which can be referenced in making educational decisions involving use of UGPA and the GRE scores.

#### Limitations of the Study

Two types of limitations need to be addressed for interpretation of the findings of this study. First, to achieve the objectives of the present study, the sample was selected by purposive sampling method, and it is not representative of the national graduate student population. The sample consisted of 28% White, 14% International students, 57% minority student for master's degree level, and 31% White, 29% International students,

39% minority students for doctoral degree level while the national graduate student population consisted of 72% White, and 28% minority (CGS, 2008). Therefore, the results may not be generalizable to other graduate schools with different demographics.

Second, findings are based on student records of a public university located in Southeast Florida during the academic year of 2000 to 2010. The characteristics of a university, location, and time period on which the present study is based limit the generalizability of the findings. The findings may have been different if the research had conducted at different times and locations.

#### Conclusions

The purpose of the study was to determine the degree of relationships among GRE scores, undergraduate GPA (UGPA), and success in graduate school, as measured by first year graduate GPA (FGPA), cumulative graduate GPA, and degree attainment status. A second aim of the study was to determine whether the relationships between the composite predictor (GRE scores, and UGPA) and the three success measures differed by race/ethnicity and sex.

The results of the correlational analyses differed by degree level. For master's students, ETS proposed prediction that GRE scores are valid predictors of first year graduate GPA for all students was supported by the findings from the present study. For master's students, first year graduate GPA was significantly and positively associated with the GRE General Test scores regardless of whether the GRE scores were examined individually or in total. However, for doctoral students, the ETS proposed prediction that GRE scores are valid predictors of first year graduate GPA for all students was partially supported by the findings from the present study.

significantly and positively associated with GREV and GRET scores, they were not significantly associated with GREQ although the correlation was positive.

The results of the regression and correlational analyses indicated that UGPA was significantly and positively associated with all three success measures (FGPA, CGPA, and degree attainment), and consistently contributed to predicting all three success measures for both master's and doctoral students.

The results of the hierarchical multiple linear and logistic regression analyses indicated that the relationships between the three success measures and the composite predictor were not significantly different among African American, Hispanic, and White students or between men and women for either degree level. However, for master's degree level, two types of the relationships, the relationship between degree attainment and GREQ for Asian Americans, and the relationship between degree attainment and UGPA for International students, were significantly different from those of White students: White students with higher GREQ scores were more likely to attain a degree within a given time limit than Asian Americans, and International students with higher UGPA were more likely to attain a degree within a given time limit than White students.

#### REFERENCES

- Ackerman, P. L. (2000). Aptitude tests. In A. E. Kazdin et al. (Eds.). *Encyclopedia of psychology* (pp. 222-224). Washington, D.C: American Psychological Association.
- Ackoff, R. L. (1965). *The design of social research*. Chicago, IL: The University of Chicago Press.
- Agresti, A. (1996). *An introduction to categorical data analysis*. New York, NY: John Wiley & Sons.
- American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- Anastasi, A. (1986). Evolving concepts of test validation. *Annual Review of Psychology*, 37, 1-15.
- Anastasi, A. & Urbina, S. (1997). *Psychological testing* (7<sup>th</sup> ed.). Upper Saddle River, NJ: Prence-Hall.
- Awad, G. H. (2007). The role of racial identity, academic self-concept, and self-esteem in the prediction of academic outcomes for African American Students. *Journal of Black Psychology*, 33(2), 188-207.
- Babbie, E. (2001). *The practice of social research* (9<sup>th</sup> ed.). Belmont, CA: Wadsworth/Thomson Learning.
- Bell, N. E. (2008). Graduate enrollment and degrees: 1997 to 2007. Washington, DC: Council of Graduate Schools. Retrieved from http://www.cgsnet.org/portals/0/pdf/R\_ED2007.pdf
- Breland, H. M. (1979). *Population validity and college entrance measures* [Research Monograph, 8]. New York, NY: The College Board.
- Brown, F. G. (1983). *Principles of educational and psychological testing* (3<sup>rd</sup> ed.). New York, NY: Holt, Rinehart and Winston.
- Brown, S. V., Clewell, B. C, Ekstrom, R. B., Goertz, M. E., & Powers, D. E. (1994). Research agenda for the Graduate Record Examinations Board minority graduate education project: An update (GRE Board Report No. MGE-94-01). Princeton, NJ: Educational Testing Service

- Burton, N. W., & Wang, M. M. (2005). Predicting long-term success in graduate school: A collaborative validity study (GRE Board Report No. 99-14R). Princeton, NJ: Educational Testing Service
- Burton, N. W., Welsh, C., Kostin, I., & VanEssen, T. (2009). Toward the definition of verbal reasoning in higher education (ETS RR-09-33). Princeton, NJ: Educational Testing Service. Retrieved from http://www.ets.org/Media/Research/pdf/RR-09-33.pdf
- Camara, W. J., & Kimmel, E. W. (2005). *Choosing students: higher education admissions tools for the 21<sup>st</sup> century*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Camp, J., & Clawson, T. (1979). The relationship between the Graduate Record Examinations Admission aptitude test and graduate grade point average in a master of arts in counseling program. *Educational and Psychological Measurement*, 39(2), 429-431.
- Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt, W. C. Borman, & Associates, *Personnel selection in* organizations (pp. 35-70). San Francisco, CA: Jossey-Bass.
- Case, D. O., & Richardson, J. V. (1990). Predictors of student performance with emphasis on gender and ethnic determinants. *Journal of Education for Library and Information Science*, *30*(3), 163-182.
- Chernyshenko, O. S., & Ones, D. S. (1999). How selective are psychology graduate programs? The effect of the selection ratio on GRE score validity. *Educational and Psychological Measurement*, *59*, 951-961.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3<sup>rd</sup> ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Council of Graduate Schools. (2000). *A walk through graduate eduation*. Washington, DC: Author.
- Council of Graduate Schools. (2008). *Graduate enrollment and degrees*: 1997 to 2007. Washington, DC: Author. Retrieved from http://www.cgsnet.org/portals/0/pdf/R\_ED2007.pdf

- Coward, W. M, & Sackett, P. R. (1990). Linearity of ability-performance relationships: A reconfirmation. *Journal of Applied Psychology*, *75*(3), 297-300.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York, NY: CBS College Publishing.
- Cronbach, L. J. (1970). *Essentials of psychological testing* (3<sup>rd</sup> ed.). New York, NY: Harper & Row.
- Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2<sup>nd</sup> ed., pp. 443-507). Washington, D.C: American Council on Education.
- Dawes, R. M. (1971). A case study of graduate admissions: Application of three principles of human decision making. *American Psychologist*, *26*(2), 180-188.
- Dawes, R. M. (1975). Graduate admission variables and future success. *Science*, 187, 721-723.
- Dobson, P. (1988). The correction of correlation coefficients for restriction of range when restriction results from the truncation of a normally distributed. *The British Journal of Mathematical & Statistical Psychology*, *41*, 227-234.
- Dollinger, S. J. (1989). Predictive validity of the Graduate Record Examination in a clinical psychology program. *Professional Psychology: Research and Practice*, 20(1), 56-58.
- Echhoff, C. M. (1966). Predicting graduate success at Winona State College. *Educational* and *Psychological Measurement*, 26(2), 483-485.
- Educational Testing Service. (2008). Graduate Record Examinations (GRE): Factors that can influence performance on the GRE general test, 2006-2007. Princeton, NJ: Author. Retrieved from http://www.ets.org/gre/edupubs
- Educational Testing Service. (2008). *GRE guide to the use of scores 2008-2009*. Princeton, NJ: Author. Retrieved from http://www.ets.org/gre/edupubs.
- Educational Testing Service. (2009). *GRE guide to the use of scores 2009-2010*. Princeton, NJ: Author. Retrieved from http://www.ets.org/gre/institutions/scores/index.html

- Educational Testing Service. (2012). *GRE guide to the use of scores 2012-2013*. Princeton, NJ: Author. Retrieved from http://www.ets.org/gre/institutions/scores/interpret/
- Educational Testing Service. (2012). *About the GRE revised General Test.* Princeton, NJ: Author. Retrieved from http://www.ets.org/gre/institutions/scores/index.html
- Enright, M. K., & Gitomer, D. H. (1989). *Toward a description of successful graduate students* (GRE Board Report No. 85-17R). Princeton, NJ: Educational Testing Service.
- Ezekiel, M., & Fox, K. A. (1959). *Methods of correlation and regression analysis: Linear and curvinear* (3<sup>rd</sup> ed.). New York, NY: John Wiley & Sons.
- Field, A. (2000). *Discovering statistics using SPSS for Windows*. London: Sage Publications.
- Florida Department of Education (2005). *State University System of Florida: Strategic plan 2005-2013*. Retrieved from http://www.flbog.org/about/\_doc/strategicplan/StrategicPlan\_05-13.pdf
- Gagne, R. M., & Medsker, K. L. (1996). *The conditions of learning: Training applications*. Orlando, FL: Harcourt Brace & Company.
- Gagne, R. M. (1962). The acquisition of knowledge. *Psychological Review*, 69(4), 355-365.
- Gay, L. R. (1996). *Educational research: Competencies for analysis and application* (5<sup>th</sup> ed.). Upper Saddle River, New Jersey: Prentice-Hall. A Simon & Schuster.
- Goldberg, E. L., & Alliger, G. M. (1992). Assessing the validity of the GRE for students in psychology: a validity generalization approach. *Educational and Psychological Measurement*, 52(4), 1019-1027.
- Green, S. B., & Salkind, N. J. (2008). Using SPSS for windows and Macintosh: analyzing and understanding the data (5<sup>th</sup> ed.). Upper Saddle River, NJ. : Pearson Prentice Hall.

Gulliksen, H. (1967). Theory of mental tests. Hillsdale, N.J: L. Erlbaum Associates.

Hackman, J. R., Wiggins, N., & Bass, A. R. (1970). Prediction of long-term success in doctoral work in psychology. *Educational and Psychological Measurement*, 30, 365-374.

- Hartnett, R. T., & Willingham, W. W. (1980). The criterion problem: What measure of success in graduate education? *Applied Psychological Measurement*, *4*, 281-291.
- Hebert, D. J., & Holmes, A. F. (1979). Graduate Record Examinations Admission aptitude test scores as a predictor of graduate grade point average. *Educational* and Psychological Measurement, 39(2), 415-420.
- Holmes, C. B., & Beishine, M. J. (1996). Correct classification, false positives, and false negatives in predicting completion of the Ph. D. from GRE scores. *Psychological Reports*, 79, 939-945.
- House, J. D. (1993). Graduate Record Examination scores and academic background variables as predictors of graduate degree completion. *Educational and Psychological Measurement*, 53, 551-556.
- House, J. D. (1997). Predictive validity of graduate record examination scores for outcomes of American Indian/Alaska Native students. *Psychological Reports*, 81(1), 337-338.
- Huck, S. W. (2004). *Reading statistics and research* (4<sup>th</sup> ed.). Boston, MA: Allyn and Bacon
- Ingram, R. E. (1983). The GRE in the graduate admission process: Is how it is used justified by the evidence of its validity? *Professional Psychology: Research and Practice, 14*(6), 711-714.
- Jaccard, J. (2001). *Interaction effects in logistic regression*. Thousand Oaks, CA: Sage publications.
- Kane, M. T. (2002). Validating high-stakes testing programs. *Educational Measurement: Issues and Practice, 21*(1), 31-41.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4<sup>th</sup> ed., pp. 17-64). Westport, CT: American Council on Measurement in Education & American Council on Education.
- Kaczmarek, M. & Franco, J. N. (1986). Sex differences in prediction of academic performance by the Graduate Record Exam. *Psychological Reports*, 59, 1197-1198.
- Keppel, G., & Zedeck, S. (1989). Data analysis for research designs: Analysis of variance and multiple regression/correlation approaches. New York, NY: W. H. Freeman and company.

- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, 127(1), 162-181.
- Linn, R. L. (2001). A century of standardized testing: Controversies and pendulum swings. *Educational Assessment*, 7(1), 29-39.
- Linn, R. L. (Ed.). (1989). *Educational measurement* (3<sup>rd</sup> ed.). American Council on Education. New York, NY: Macmillan Publishing Company.
- Longford, N. T. (1991). *Negative Coefficients in the GRE Validity Study Service* (GRE Board Report No. 89-05P). Princeton, NJ: Educational Testing Service.
- Madaus, G. F., & Walsh, J. J. (1965). Departmental differentials in the predictive validity of the Graduate Record Examination Aptitude Tests. *Educational and Psychological Meausrement, 25*(4), 1105-1110.
- McCloy, R. A., Campbell, J. P., & Cudeck, R. (1994). A confirmatory test of a model of performance determinants. *Journal of Applied Psychology*, *79*(4), 493-505.
- Mehrens, W. A., & Lehmann, I. J. (1987). Using standardized tests in education (4<sup>th</sup> ed.). New York, NY: Longman.
- Menard, S. W. (1995). *Applied logistic regression analysis*. Thousand Oaks, CA: SAGE Publications.
- Menard, S. (2002). *Applied logistic regression* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage Publications.
- Merriam, S. B., & Simpson, E. L. (1995). *A guide to research for educators and trainers of adults* (2<sup>nd</sup> ed.). Malaba, FL: KRIEGER.
- *Merriam-Webster's collegiate dictionary* (11<sup>th</sup> ed.). (2003). Springfield, MA: Merriam-Webster.
- Mertler, C. A., & Vannatta, R. A. (2010). *Advanced and multivariate statistical methods: Practical application and interpretation* (4<sup>th</sup> ed.). Glendale, CA: Pyrczak Publishing.
- Moore, D. S., & McCabe, G. P. (1998). *Introduction to the practice of statistics* (3<sup>rd</sup> ed.). New York, NY: W. H. Freeman and Company.
- Morrison, T., & Morrison, M. (1995). A meta-analytic assessment of the predictive validity of the quantitative and verbal components of the Graduate Record

Examination with graduate grade point average representing the criterion of graduate success. *Educational and Psychological Measurement*, 55(2), 309-316.

- National Center for Education Statistics (2007). *The path through graduate school: A longitudinal examination 10 years after bachelor's degree*. Retrieved from http://nces.ed.gov/pubs2007/2007162.pdf
- National Center for Education Statistics (2008). *Total number of degree-granting institutions and fall enrollment in those institutions, by type and control of institution and percentage of minority enrollment: 2007.* Retrieved from http://nces.ed.gov/programs/digest/d08/tables/dt08\_230.asp
- National Center for Education Statistics. (2009). *The condition of graduate education 2008: Trends in graduate and first professional enrollments*. Retrieved from http://nces.ed.gov/programs/coe/2008/section1/table.asp?tableID=872
- National Center for Education Statistics (2012). Conditions of education 2012: Note 1: Commonly used measures. Race/Ethnicity. http://nces.ed.gov/programs/coe/supnotes/2012-n01.asp
- National Center for Education Statistics. (2012). *Status and trends in the education of racial and ethnic minorities*. Retrieved from http://nces.ed.gov/pubs2010/2010015/index.asp
- National Center for Education Statistics. (2012). *Integrated postsecondary education data system: Definitions for new race and ethnicity categories*. Retrieved from http://nces.ed.gov/ipeds/reic/definitions.asp
- Nunnally, J. C. (1978). *Psychometric theory* (2<sup>nd</sup> ed.). New York, NY: McGraw-Hill Book.
- Oldfield, K., Hutchinson, J. R. (1997). Predictive validity of the Graduate Record Examination with and without range restrictions. *Psychological Reports*, 81(1), 211-220.
- Omizo, M. M., & Michael, W. B. (1979). The prediction of performance in a counselor education master's degree program. *Educational and Psychological Measurement*, 39(2), 433-437.
- Ozer, D. J. (1985). Correlation and the coefficient of determination. *Psychological Bulletin*, *97*, 307-315.
- Panpel, F. C. (2000). Logistic regression: A primer. Thousand Oaks, CA: Sage.

- Pressley, M., & McCormick, C. B. (1995). Advanced educational psychology for educators, researchers, and policymakers. New York, NY: HarperCollins College.
- Ruscio, J. (1998). Gatekeeping, compensation, and fallibility. *American Psychologist*, 53(5), 568-569.
- Sacks, P. (2001, June 8). How admissions tests hinder access to professional and graduate school. *Chronicle of Higher Education*, B11-B12.
- Sampson, C., & Boyer, P. G. (2001). GRE scores as predictors of minority students' success in graduate study: An argument for change. *College Student Journal*, 35(2), 271-279.
- Schmidt, F. L. & Hunter, J. E. (1981). Employment Testing: Old theories and new research findings. *American Psychologist*, 36(10), 1128-1137.
- Schneider, L. M., & Briel, J. B. (1990). Validity of the GRE: 1988-1989 summary report (GRE Board Report No. 90-01VSS). Princeton, NJ: Educational Testing Service.
- Scott, R. R., & Shaw, M. E. (1985). Black and White performance in graduate school and policy implications of the use of graduate record examination scores in admissions. *Journal of Negro Education*, 54 (1), 14-23.
- Singleton, R. Jr., Straits, B. C., Straits, M. M., & McAllister, R. J. (1988). Approaches to social research. New York, NY: Oxford University Press.
- Smallwood, S. (2004, January 16). Doctoral dropout. *The Chronicle of Higher Education*, 50(19), A10-A12.
- Sternberg, R. J., & Williams, W. M. (1997). Does the Graduate Record Examination predict meaningful success in the graduate training of psychologists? *American Psychologist*, 52(6), 630-641.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3<sup>rd</sup> ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Stricker, G., & Huber, J. T. (1967). The Graduate Record Examinations and undergraduate grades as predictors of success in graduate school. *The Journal Educational Research*, 60, 466-468.
- Sweet, S. A., & Grace-Martin, K. (2003). *Data analysis with SPSS: A first course in applied statistics* (2<sup>nd</sup> ed.). Boston, MA: Pearson Education, Inc.
- Tabachnick, B. G., & Fidell, L. S. (1983). *Using multivariate statistics*. Cambridge, NY: Harper & Row.

- Thacker, A. J., & Williams, R. E. (1974). The relationship of the Graduate Record Examination to grade point average success in graduate school. *Psychological and Educational Measurement*, *34*, 939-944.
- UCLA: Academic Technology Services, Statistical Consulting Group. (2011). *How do I interpret odds ratio in logistic regression?* Retrieved from http://www.ats.ucla.edu/stat/mult\_pkg/faq/general/odds\_ratio.htm
- Thompson, B. (2002). "Statistical," "Practical," and "Clinical": How many kinds of significance do counselors need to consider? *Journal of Counseling & Development*, 80, 64 71.
- U.S. Department of Education, Institute of Education Sciences NCES. (2007). *Status and Trends in the Education of Racial and Ethnic Minorities* (No. NCES 2007-039). Washington, DC.
- Wainer, H., & Braun, H. I. (Eds). (1988). *Test validity*. Hillsdale, N.J: LAWRENCE ERLBAUM ASSOCIATES.
- Walpole, M., Burton, N. W., Kanyi, K., & Jackenthal, A. (2002). Selecting successful graduate students: In-depth interviews with GRE users (GRE Board Report No. 99-11R). Princeton, NJ: Educational Testing Service.
- Whitworth, R. H., & Barrientos, G. A. (1990). Comparison of Hispanic and Anglo graduate record examination scores and academic performance. *Journal of Psychoeducational Assessment*, 8(2), 128-132.
- Willingham, W. W. (1974). Predicting success in graduation education. *Science*, 187, 273-278.
- Willingham, W. W. (1976). Validity and the Graduate Record Examinations Program (GRE Board Report No. 76-01SR). Princeton, NJ: Educational Testing Service.
- Wilson, K. M. (1979). The validation of GRE scores as predictors of first-year performance in graduate study: Report of the GRE cooperative validity studies project (GRE Board Report No. 75-08R). Princeton, NJ: Educational Testing Service.
- Young, J. W. (2001). Differential validity, differential prediction, and college admission testing: A comprehensive review and analysis (College Board Research Report No. 2001-6). Princeton, NJ: Educational Testing Service.
- Zwick, R., & Braun, H. I. (1991). An analysis of graduate school careers in three universities: Differences in attainment patterns across academic programs and

*demographic groups* (GRE Board Report No. 86-21p). Princeton, NJ: Educational Testing Service.

Zwick, R. (2002). Fair game? The use of standardized admissions tests in higher education. New York, NY: RoutledgeFalmer.

APPENDICES

			GREV			GREQ			GRET			UGPA	
Characteri	stic	Ν	М	SD	N	М	SD	N	M	SD	N	М	SD
Sex	Women	3,722	431.56	97.34	3,722	497.38	129.06	3,722	928.80	189.35	3,591	3.38	.40
	Men	2,268	435.38	106.67	2,268	585.26	135.10	2,268	1,020.55	188.71	2,187	3.25	.44
Race/	African Am.	789	406.79	90.85	789	459.65	127.41	789	866.13	181.13	755	3.17	.44
Ethnicity	Asian Am.	229	425.85	106.31	229	583.28	129.33	229	1,008.95	180.41	224	3.26	.44
	Hispanic	2,369	417.06	91.29	2,369	496.47	129.53	2,369	913.40	183.86	2,279	3.36	.40
	Int'l	829	405.22	110.37	829	647.25	131.05	829	1,052.46	183.75	804	3.35	.40
	Native Am.	10	447.00	92.98	10	559.00	128.88	10	1,006.00	198.06	10	3.21	.38
	NR	70	418.00	86.84	70	581.57	114.69	70	999.57	149.38	66	3.49	.32
	White	1 694	482.61	95 98	1 694	545.08	117.63	1 694	1 027 63	177 87	1 640	3 36	43

Descriptive Statistics for Master's Students' GRE Scores and UGPA by Sex and Race/Ethnicity

White1,694482.6195.981,694545.08117.631,6941,027.63177.871,6403.36Note. GREV is GRE Verbal Reasoning Test score. GREQ, GRE Quantitative Reasoning Test score, GRET, sum of GREV and GREQ. UGPA is<br/>undergraduate GPA. Am. is abbreviation of American. Int'l is an abbreviation of International students. NR is a group whose race or ethnicity was not<br/>identified.

			FGPA			CGPA		Ι	Degree Atta	inment	
Characteristi	с	N	М	SD	N	М	SD	Ν	М	f	%
Sex	Women	3,699	3.54	.52	2,999	3.46	.75	3,317	34.49	2,839	86.0
	Men	2,249	3.43	.56	1,834	3.33	.77	1,864	38.38	1,518	81.0
Race/	African Am.	731	3.34	.57	589	3.25	.78	639	37.72	527	83.0
Ethnicity	Asian Am.	212	3.51	.57	163	3.40	.81	196	34.15	165	84.0
	Hispanic	2,146	3.52	.50	1,670	3.40	.78	1,945	34.27	1,674	86.0
	Int'l	744	3.46	.48	605	3.44	.61	514	35.11	441	86.0
	Native Am.	10	3.28	.45	7	2.77	1.54	10	24.30	9	90.0
	NR	62	3.59	.37	54	3.32	.84	55	37.32	45	82.0
	White	1,474	3.59	.55	1,245	3.49	.79	1,341	38.18	1,100	82.0

Descriptive Statistics for Master's Students' FGPA, CGPA, and Degree Attainment by Sex and Race/Ethnicity

*Note.* FGPA is first year graduate GPA. CGPA is cumulative graduate GPA. Am is an abbreviation of American. Int'l is an abbreviation of International students. NR is a group whose race or ethnicity was not identified. The mean, *M*, under Degree Attainment column is the time elapsed (in months) until degree attainment. *f* is the number of students who attained the degree.

			GREV			GREQ			GRET			UGPA	
Characteri	istic	N	М	SD	N	М	SD	N	M	SD	N	М	SD
Sex	Women	710	493.75	99.76	710	592.82	123.89	710	1,086.42	173.69	687	3.56	.36
	Men	667	488.20	115.18	667	666.52	116.37	667	1,154.69	171.78	638	3.48	.39
Race/	African Am.	135	466.37	89.47	135	532.22	122.35	135	998.59	173.24	132	3.43	.44
Ethnicity	Asian Am.	47	473.83	11503	47	638.30	120.93	47	1,112.13	171.76	46	3.40	.38
	Hispanic	327	466.64	93.01	327	572.87	125.91	327	1,039.27	172.32	314	3.50	.40
	Int'l	405	471.21	1,21.95	405	722.00	88.00	405	1,193.21	159.82	388	3.51	.34
	Native Am.	6	571.67	48.34	6	635.00	57.53	6	1,206.67	33.27	6	3.58	.47
	NR	24	467.92	118.32	24	650.00	100.48	24	1,117.92	156.40	22	3.56	.39
	White	433	537.81	91.41	433	610.79	103.07	433	1.148.50	152.20	417	3.60	.36

Descriptive Statistics for Doctoral Students' GRE Scores and UGPA by Sex and Race/Ethnicity

*Note.* GREV is GRE Verbal Reasoning Test score. GREQ, GRE Quantitative Reasoning Test score, GRET, sum of GREV and GREQ. UGPA is undergraduate GPA. Am. is abbreviation of American. Int'l is an abbreviation of International students. NR is a group whose race or ethnicity was not identified.

			FGPA			CGPA			Degree Att	ainment	
Characteristi	c	N	М	SD	N	М	SD	N	М	f	%
Sex	Women	710	3.68	.37	551	3.67	.40	244	55.29	187	77.0
	Men	667	3.62	.49	508	3.59	.56	202	53.11	151	75.0
Race/	African Am.	135	3.57	.44	104	3.55	.42	49	91.67	29	59.2
Ethnicity	Asian Am.	47	3.53	.54	37	3.59	.54	8	58.75	8	100.0
	Hispanic	327	3.66	.35	268	3.68	.32	112	86.16	84	75.0
	Int'l	405	3.65	.39	292	3.64	.48	101	65.27	82	81.2
	Native Am.	6	3.47	.76	3	3.64	.33	2	66.00	2	100.0
	NR	24	3.48	.66	21	3.51	.62	12	86.50	7	58.3
	White	433	3.68	.48	334	3.63	.60	162	76.52	126	77.8

Descriptive Statistics for Doctoral Students' FGPA, CGPA, and Degree Attainment by Sex and Race/Ethnicity

*Note.* FGPA is first year graduate GPA. CGPA is cumulative graduate GPA. Am is an abbreviation of American. Int'l is an abbreviation of International students. NR is a group whose race or ethnicity was not identified. The mean, M, under Degree Attainment column is the time elapsed (in months) until degree attainment. f is the number of students who attained the degree.

Means, Standard Deviations, and Zero-Order Correlation Coefficients of GRE Scores and UGPA with FGPA by Sex and Race/Ethnicity

				Master	r's						Doctor	al		
		FC	<b>BPA</b>						FC	GPA				
Characteristic	N	М	SD	GREV	GREQ	GRET	UGPA	N	М	SD	GREV	GREQ	GRET	UGPA
Sex														
Women	3,371	3.55	0.51	.17**	.10**	.15**	.22**	710	3.68	0.37	.11**	.11**	.14**	.16**
Men	2,008	3.44	0.56	.14**	.08**	.13**	.19**	667	3.62	0.49	.04	.02	.04	.21**
Race/Ethnicity														
African Am.	731	3.34	0.57	.08*	02	.03	.16**	135	3.57	0.44	.07	.17	.15	.19*
Asian Am.	212	3.50	0.57	.15*	05	.05	.26**	47	3.53	0.54				
Hispanic	2,146	3.52	0.50	.10**	.05**	.09**	.23**	327	3.66	0.35	.14*	03	.05	.27**
International	744	3.46	0.48	.13**	.10**	.15**	.17**	405	3.65	0.39	01	.01	001	.14**
Native Am.	10	3.28	0.45					6	3.47	0.76				
Not Reported	62	3.59	0.37	.24*	.30*	.37**	.51**	24	3.48	0.66				
White	1,474	3.59	0.55	.16**	.03	.11**	.18**	433	3.68	0.48	.13**	.02	.09	.19**

*Note.* FGPA is an abbreviation of first year graduate GPA. GREV is GRE Verbal Reasoning Test scores. GREQ is GRE Quantitative Reasoning Test score. GRET is the sum of GREV and GREQ scores. UGPA is undergraduate GPA. Am is an abbreviation of American. Int'l is an abbreviation of International students. NR is a group whose race or ethnicity was not identified. Due to the small sample sizes, Native Americans were excluded from the data analyses for both master's degree and doctoral degree level analyses; Asian Americans and NRs were additionally excluded from the doctoral degree level analyses for the same reason.

\* *p* < .05. \*\* *p* < .01.

Means, Standard Deviations, and Zero-Order Correlation Coefficients of GRE Scores and UGPA with CGPA by Sex and Race/Ethnicity

				Maste	r's							Doctor	al		
		CO	GPA							CGF	PA				
Characteristic	N	М	SD	GREV	GREQ	GRET	UGPA	N		М	SD	GREV	GREQ	GRET	UGPA
Sex															
Women	2,704	3.46	0.76	.11**	.01**	.12**	.15**	55	1	3.67	0.40	.07	.09*	.10*	.13**
Men	1,629	3.33	0.77	.08**	.07**	.09**	.17**	50	8	3.59	0.56	.05	.07	.08	.12**
Race/Ethnicity															
African Am.	589	3.25	0.78	.07	.06	.08	.15**	10	4	3.55	0.42	04	.14	.08	.28**
Asian Am.	163	3.40	0.81	.09	01	.05	.13	3	7	3.59	0.54				
Hispanic	1,670	3.40	0.78	.08**	.03	.06*	.16**	26	8	3.68	0.32	.21**	.08	.17**	.23**
Int'l	605	3.44	0.61	.13**	.14**	.18**	.18**	29	2	3.64	0.48	03	.06	.01	.03
Native Am.	7	2.77	1.54	.14	04	.05	.09		3	3.64	0.33				
NR	54	3.32	0.84					2	1	3.51	0.62				
White	1,245	3.49	0.79	.07**	.02	.05*	.16**	33	4	3.63	0.60	.14*	.02	.01	.15**

*Note.* CGPA is an abbreviation of cumulative graduate GPA. GREV is GRE Verbal Reasoning Test scores. GREQ is GRE Quantitative Reasoning Test score. GRET is the sum of GREV and GREQ scores. UGPA is undergraduate GPA. Am is an abbreviation of American. Int'l is an abbreviation of International students. NR is a group whose race or ethnicity was not identified. Due to the small sample sizes, Native Americans were excluded from the data analyses for both master's degree and doctoral degree level analyses; Asian Americans and NRs were additionally excluded from the doctoral degree level analyses for the same reason.

\* *p* < .05. \*\* *p* < .01.

				Master's							Doctora	ıl		
		Degre	ee						Degree	e				
Characteristic	Ν	M	SD	GREV	GREQ	GRET	UGPA	Ν	М	S	GREV	GREQ	GRET	UGPA
										D				
Sex														
Women	3,392	34.49		10**	02	07**	.05**	710	55.29		.01	.14*	.11	.08
Men	2,023	38.38		08**	.06**	001	.06**	667	53.11		.08	.17*	.16*	.15*
Race/Ethnicity														
African Am.	740	37.72		11**	02	07	.08*	135	61.20		.20	.28	.31*	.29
Asian Am.	212	34.15		01	16*	17*	.04	47	49.77					
Hispanic	2,161	34.27		09**	.03	02	.06**	327	61.39		.01	.05	.04	.05
Int'l	747	35.11		02	.03	.01	.20**	405	44.63		.03	05	01	04
Native Am.	10	24.30						6	41.33					
NR	62	37.32		.23	11	.06	.10	24	73.96		11	.34	.16	10
White	1,483	38.18		12**	02	08**	.01	433	55.21		.02	.15	.12	.18*

Means, Standard Deviations, and Zero-Order Correlation Coefficients of GRE Scores and UGPA with Degree Attainment within a Given Time Limit by Sex and Race/Ethnicity

*Note.* Degree is an abbreviation of degree attainment. GREV is GRE Verbal Reasoning Test scores. GREQ is GRE Quantitative Reasoning Test score. GRET is the sum of GREV and GREQ scores. UGPA is undergraduate GPA. Am is an abbreviation of American. Int'l is an abbreviation of International students. NR is a group whose race or ethnicity was not identified.NR is a group whose race or ethnicity was not identified. M in line with Characteristic under Degree is the time elapsed (in months) until degree attainment. Due to the small sample sizes, Native Americans were excluded from the data analyses for both master's degree and doctoral degree level analyses; Asian Americans and NRs were additionally excluded from the doctoral degree level analyses for the same reason.

\* *p* < .05. \*\* *p* < .01.

Summary of Relationships between GREV, GREQ UGPA, and FGPA, CGPA, Degree Attainment Status for Research Question 1, 2, and 3(Master's Degree Level)

				G	RE Scores					UGPA		Composite
	N	V	V	V	Q	Q	Q	Т	U	U	U	V+Q+U
		(r)	$(\beta)$	(OR)	(r)	(β)	(OR)	(r)	(r)	$(\beta)$	(OR)	$R^2$ (adj. $R^2$ )
												.06 (.06)
FGPA	5,738	.15**	.13***		.05**	.02		.12**	.21**	.20***		V+U
												.04 (.04)
CGPA	4,654	.10**	.07***		.06**	.04*		.09**	.17**	.16***		V+Q+U
								05**				.02; .03
Degree	5,015	10**		.10***	00		1.001*		.06**		1.56***	Q+U

*Note.* V is an abbreviation of GRE Verbal Reasoning Test score; Q, GRE Quantitative Reasoning Test score; U, undergraduate GPA. T is the total scores of GRE Verbal and GRE Quantitative Reasoning test scores. FGPA is first year graduate GPA. CGPA is cumulative graduate GPA. Degree is abbreviation of degree attainment status. The indices in the column, Composite, are regression coefficients yielded from the regression analyses of each success measure on the composite score of GRE Verbal, GRE Quantitative, and undergraduate GPA. Adj. is an abbreviation of adjusted  $R^2$ . \*p < .05. \*\*p < .01. \*\*\* p < .001.

Summary of Relationships between GREV, GREQ UGPA, and FGPA, CGPA, Degree Attainment Status for Research Question 1, 2, and 3 (Doctoral Degree Level)

				(	GRE Score	S				UGPA		Composite
	N	V	V	V	Q	Q	Q	Т	U	U	U	V+Q+U
		(r)	$(\beta)$	(OR)	(r)	(β)	(OR)	(r)	(r)	(β)	(OR)	$R^2$ (adj. $R^2$ )
												.04 (.04)
FGPA	1,325	.07**	.05		.03	.04		.07*	.19**	.09***		Ŭ
												.02 (.02)
CGPA	1,013	.06	.04		.05	.05		.07*	.13**	.13***		Ù
												.03; .04
Degree	400	.05		1.00	.14**		1.002**	.13**	.12*		1.92*	O+U

*Note.* V is an abbreviation of GRE Verbal Reasoning Test score; Q, GRE Quantitative Reasoning Test score; U, undergraduate GPA. T is the total scores of GRE Verbal and GRE Quantitative Reasoning test scores. FGPA is first year graduate GPA. CGPA is cumulative graduate GPA. Degree is abbreviation of degree attainment status. The indices in the column, Composite, are regression coefficients yielded from the regression analyses of each success measure on the composite score of GRE Verbal, GRE Quantitative, and undergraduate GPA. Adj. is an abbreviation of adjusted  $R^2$ . \*p < .05. \*\*p < .01. \*\*\*p < .001.

Summary of Interactions of Race/Ethnicity by the Composite Predictor on Graduate Students' FGPA, CGPA, and Degree Attainment Status for Research Question 4

			Master	''s				Doctoral		
	Ν	V	Q	U	V+Q+U	Ν	V	Q	U	V+Q+U
FGPA	5,738				NO	1,325		NR		YES
CGPA	4,654				NO	1,013				NO
Degree	5,015		Asian	Int'l	YES	400				NO

*Note*. Master's represents master's degree level. Doctoral represents doctoral degree level. V is an abbreviation of GRE Verbal Reasoning Test score; Q, GRE Quantitative Reasoning Test score; U, undergraduate GPA. FGPA is first year graduate GPA. CGPA is cumulative graduate GPA. Degree is abbreviation of degree attainment status. Asian is an abbreviation of Asian American students.

NO represents no significant interaction. YES represents significant interaction. NR, in line with FGPA under column Q in Doctoral degree level represents that the interaction of NR by GREQ on doctoral student's FGPA was significant. Asian, in line with Degree under column Q in master's degree level represents that the interaction GREQ by Asian Americans on master's degree attainment was significant. Int'l in line with Degree under column U in master's degree level represents that the interaction of UGPA by International student on master's degree attainment was significant.

*Validity Data of the GRE Scores and UGPA for Success in Graduate School from Meta – Analytic Studies (White Majority)* 

Author	Data	# of						
Year	from	Studies	N: Total	Criteria of		Predicto	r <sup>a</sup>	
(Affiliation)	Period	Reviewed	(Level)	Success Measures	V	Q	Т	U
Thacker & Williams <sup>1</sup>	1957~1970	12	569 (N/S)	1st semester grad GPA	.18*	.19*		
1974 (Independent)			1,154(N/S)	9, 12, 15 Hours GGPA	.21*to.37	.08to.36	.18to.34	
			42(N/S)	First year Grad GPA	.22	.15		
			252(N/S)	Graduated/not-grad	.28*to.34**	.08to.21*		
			252	Doctoral GPA	.21*	.32**	.31	
			24	Master's GPA	.49*	.37	.54**	
Willingham, W. W. <sup>2</sup>	1952~1972	43	21,214	Graduate GPA	.24(46 <sup>b</sup> )	.23 (43 <sup>b</sup> )		.31(26 <sup>b</sup> )
1974 (ETS)			$(N/S^{c})$	Attain Ph. D.	.18(47 <sup>b</sup> )	.26 (47 <sup>b</sup> )		.14(30 <sup>b</sup> )
Wilson, K. M. <sup>3</sup>	1974~1975	39 <sup>d</sup>	4,433	First year Grad GPA	.05to.43	.04to.52		.06to.56
1979 (ETS)			M &D					
Schneider & Briel <sup>4</sup>	1983~1988	606	9,200	First year Grad GPA	.29	.28		0.34
1990 (ETS)			(N/S)	Faculty rating	.25	.25		0.31
Goldberg & Alliger <sup>5</sup>	1950~1990	27	2,754	GGPA	.15	.15		
1992 (Independent)			(N/S)	Comprehensive Exam.	.37	.28		
Morrison & Morrison	1955~1992	22	5,186	GGPA	.28	.22		
1995 (Independent)			(N/S)					
Kuncel, Hazlett, & Ones <sup>6</sup>	1940s~1990s	1,753	82,659	First year Grad GPA	.24 (.34 <sup>e</sup> )	.24 (.38)		.30 (.33)
2001 (Independent)			(N/S)	Graduate GPA	.23 (.34)	.21 (.32)		.28 (.30)
				Degree attainment	.14 (.18)	.14 (.20)		.12 (.12)

*Note.* Adapted from "The relationship of the Graduate Record Examination to Grade Point Average and Success in graduate school" by A. J. Thacker, and R. E.Williams, 1974, Educational and Psychological Measurement, 34, p. 941. Copyright 1974 by Educational and Psychological Measurement. <sup>a</sup>Predictor: V: GRE Verbal Reasoning Test score; Q: GRE Quantitative Reasoning scores. T: total of V and Q. U: Undergraduate GPA. <sup>b</sup>The number of studies on which the median *r* is based. <sup>c</sup>N/S: not specified. <sup>d</sup>Number of schools participated in the study. <sup>e</sup>The coefficients corrected for the range restriction and criterion unreliability.

\**p* < .05. \*\**p* < .01.

#### Footnotes

<sup>1</sup> Thacker and Williams (1974). Criteria in the table are selected from the study only when the relationship includes at least one significant predictor-criterion relationship. Hence the total in the table does not equal to the total number of subjects (1,874) of the data set. Sample size within the study: Because the data are pooled from 12 studies, the number of subjects used for each criterion varies. Validity coefficients of the GRE scores for various criterion measures are summarized in ranges to show the variation of the each predictor. Degree levels are not specified except when it is obvious from the name of the criterion measures.

<sup>2</sup> Willingham (1974). The author indicated that 43 studies included 138 independent sets of data, usually corresponding to departments with some exceptions. The individual sets of data comprised 20 to 1,479 student records (median N = 80). The numbers entered for validity coefficients are median *r* between various predictors and criteria of success in graduate school.

<sup>3</sup> Wilson (1979). The validity coefficients entered in the table are the ranges of the values summarized from the field/department data. The validity coefficient of .06 involving UGPA and first year GGPA was based on only one sample. By this study, ETS established the reporting system in which GRE validity coefficients are reported by field/department. The validity coefficients are expressed in ranges showing the variations by field/department in the table.

<sup>4</sup> Schneider and Briel (1990). The authors noted that the validity coefficients in this study are size-adjusted average correlations.

<sup>5</sup> Goldberg and Alliger (1992). The authors did not provide the total N. Instead, they indicated that the average number of subjects was 102, ranging from 23 to 582. Thus, the total N is calculated by multiplying the average of 102 by 27 (number of the studies), which amounts to 2,754.

<sup>6</sup> Kuncel, Hazlet, & Ones (2001). The authors stated that 1,753 independent samples data were included in the meta-analysis. The observed validity coefficients entered first in the cell; those in parentheses are the validity coefficients corrected for range restriction and criterion unreliability. The number of studies reviewed was not specified in the study. The information entered in the table is the number of independent samples data the study was based on.

Validity Date of the GRE Scores and UGPA for Success in Graduate School by Department from Meta - Analytic Studies (White Majority)

	Author	Data	# of						
	Year	From	studies	N: Total	criteria of		Predictor		
Sex <sup>a</sup>	(Affiliation)	Period	Reviewed	(Level <sup>b</sup> )	success measures	V	Q	Т	U
N/S	Wilson, K. M.	1974~	39	4,433	First year GGPA	.05 to .43	.04 to .52		.06 to .56
	1979 (ETS)	1975		M &D					
M: 4,498	Schneider, L. M.,	1983~	606	9,200	First year GGPA	.22 to .31	.18 to .32		.29 to .39
W: 4,700	& Briel, J. B. 1990 (ETS)	1988		(N/S)	Faculty rating	.25 to .26	.23 to .31		.30 to .34
N/S	Kuncel, N. R.,	N/S	1,753	82,659					
	Hezlett, S. A., &			(N/S)	First year GGPA	.16 to .28	.23 to .25		.30 to .31
	Ones, D. S.,				Graduate GPA	.21 to .27	.18 to .25		.13 to .38
	2001 (Independent)				Faculty ratings	.20 to .41	.20 to .34	0.46	.19 to .25
					Degree attainment	.03 to .41	07 to .22		02 to .22
M: 472;	Burton, N. W.	1995~	7	1,094	GGPA			.29 to .39	.11 to .29
W: 831	& Wang, M.M.	1998	schools	(M:639;					
	2005 (ETS)		(21 dpts <sup>c</sup> )	D664)					

*Note.* V: GRE Verbal Reasoning Test scores; Q: GRE Quantitative Reasoning Test scores; T: total of V and Q. UGPA: Undergraduate GPA. The validity data of all ETS researchers and those of Kuncel et al. are in ranges, indicating differences across different departments. Cells are left blank when the data are not available from the studies. <sup>a</sup>Sex: Men and women combined. <sup>b</sup>Level: N/S: Not specified; M: Masters; D: Doctoral. <sup>c</sup>7 schools and 21 depts represent the number of schools from which the data were collected for the study.
Regression Coefficients of the GRE Scores and UGPA for Success in Graduate School by Department from Metaanalytic Studies (White Majority)

	Author	Data	# of				Predictor	
	Year	From	studies	N: Total	criteria of	V	Q	U
Sex <sup>a</sup>	(Affiliation)	Period	Reviewed	(Level <sup>b</sup> )	success measures	(β)	$(\beta)$	$(\beta)$
N/S	Wilson, K. M. 1979, ETS	1974-1975	39	4,433 M &D	First year GGPA	08 to .37	.08 to .37	.18 to .33
M: 472; W: 831	Burton, N. W. & Wang, M. M. 2005 (ETS)	1995~1998	7 schools (21 dpts <sup>c</sup> )	1,094 (M:639; D664)	GGPA	.06 to .20	.01 to .19	.02 to .25

*Note.* V: GRE Verbal Reasoning Test scores; Q: GRE Quantitative Reasoning Test scores; UGPA: Undergraduate GPA. The validity data are in ranges, indicating differences across different departments. <sup>a</sup>Sex: Men and women combined. N/S: Not specified; <sup>b</sup>Level: M: Masters; D: Doctoral. <sup>c</sup>7 schools and 21 depts represent the number of schools and departments from which the data were collected for the study.

Author		Subjects							
Year		(Sex <sup>a</sup> )		Criterion	N: Total				
(Affiliation)	Location	Ethnicity <sup>b</sup>	Predictor <sup>c</sup>	measure	(Level <sup>d</sup> )	M	r	β	
Sampson & Boyer	N/R <sup>e</sup>	(Combined:	V	GGPA	160	447.0	0.39***	0.35	
2001		W: 103; M: 57)	Q		(M: 108; D: 50;	450.0	0.05	N/R	
(Independent)		AA: 144; H: 13;	А		S: 1)	470.0	0.11	N/R	
		AI: 2; A: 1	U			3.11	0.15	0.18	

Validity Data of the GRE Scores and UGPA for Success in Graduate School (African American Majority)

*Note.* n = 160. Cells are left blank when the data are not available from the studies.

<sup>a</sup>Sex: men and women combined. <sup>b</sup>Ethnicity: AA: African American; H: Hispanic; AI: American Indian; A: Asian American. <sup>c</sup>Predictor: V: GRE Verbal Reasoning Test scores; Q: GRE Quantitative Reasoning Test scores; A: GRE Analytical Test scores. U: Undergraduate GPA. <sup>d</sup>Level: M: Masters; D: Doctoral; S: Specialist. <sup>e</sup>N/R: Not reported.

\*\*\**p* < .01.

Author		Subjects							
Year		(Sex <sup>a</sup> )		criterion	N: Total				
(Affiliation)	Location	Ethnicity	Predictor <sup>b</sup>	measure	(Level <sup>c</sup> )	M	SD	r	R
House, J. D.	IL	(Combined)		Overall	28 (M: 26; D:2)				
1997		American	V	CGPA	26	494.6	92.80	0.02	
(Independent)		Indian	Q		26	521.4	135.50	- 0.08	
			V+Q		26	1016.6	205.50		- 0.04
			V	Degree-		494.6	92.80	0.20	
			Q	Completion	1	521.4	135.50	0.47*	
			V+ Q	-		1016.6	205.50		0.40*

# Validity Data of the GRE Scores for Success in Graduate School (American Indian)

*Note.* n = 28. Cells are left blank when the data are not available from the studies. <sup>a</sup>Sex: Men and women combined. <sup>b</sup>Predictor: V: GRE Verbal Reasoning Test scores; Q: GRE Quantitative Reasoning Test scores. <sup>c</sup> Level: M: Masters; D: Doctoral.

\**p* < .05.

Author		Subjects (Sex <sup>a</sup> )		criterion	N: Total					
	<b>.</b> .		D I h	criterion	IV. TOtal	1.6	a D	P	<b>n</b> <sup>2</sup>	0
(Affiliation)	Location	Ethnicity	Predictor	measure	(Level <sup>e</sup> )	M	SD	R	Rž	β
Whitworth, R. H. &	TX	(Combined)			952					
Barrientos, G. A.		Hispanic	V	GGPA	320 (N/S)	429.8	87.00			0.000
1990			Q			444.0	113.30			0.000
(Independent)			А			429.6	100.00			0.000
			U			3.0	0.47			0.28
			V+Q+A+U					0.19	0.04	
		(Combined)	V	GGPA	632 (N/S)	515.0	95.50			0.001
		Anglo	Q		× ,	502.8	116.40			0.000
		C	Ă			508.6	110.90			0.000
			U			3.2	0.45			0.25
			V+Q+A+U					0.27	0.07	

Validity Data of the GRE Scores and UGPA for Success in Graduate School (Hispanic and Anglo Students)

*Note.* n = 952. Cells are left blank when the data are not available from the studies. <sup>a</sup>Sex: men and women combined. <sup>b</sup>Predictor: V: GRE Verbal Reasoning Test scores; Q: GRE Quantitative Reasoning Test scores; A: GRE Subject Test scores; U: Undergraduate GPA. <sup>c</sup> Level: N/S: Not specified.

Author		Subjects							
Year		Sex		criterion	N: Total				
(Affiliation)	Location	(Ethnicity)	Predictor	measure	(Level <sup>a</sup> )	M	SD	r	R
Kaczmarek,	New								
M., &	Mexico	Overall			43 (M)				
Franco, J. N.		Men	V	Final GGPA	18	527.22	99.16	0.11	
1986		(Caucasian)	Q			525.56	77.32	- 0.04	
(Independent)			V+Q						$0.000^{b}$
		Women	V	Final GGPA	25	504.00	97.08	0.24	
		(Caucasian)	Q			461.20	93.77	0.56*	
			V+Q						0.52*

Validity Data of the GRE Scores and UGPA for Success in Graduate School (Caucasian Men and Women)

*Note*. n = 43. V: GRE Verbal Reasoning Test scores; Q: GRE Quantitative Reasoning Test scores. Cells are left blank when the data are not available. <sup>a</sup>The sample data is based on students in a terminal Master's program in Counseling. <sup>b</sup>The validity coefficient for Caucasian Men is .000, not missing data. \*p < .05.

### VITA

#### MYUNG SOOK HYUN

	Born, Seoul, South Korea
	B. As, English Language and Literature; and Buddhist Studies Dongguk University Seoul, South Korea
1983-1988	Regency Club Manager, Grand Hyatt Seoul, Korea Trainer, Grand Opening Support Team member, Hyatt International Corporation Hyatt Regency Cheju, South Korea Hyatt Regency Melbourne, Australia
1991	M. S., Hotel and Foodservice Management Florida International University Miami, Florida
1992-1997	Hotel Training Manager The Westin Chosun Hotel Seoul, South Korea
1997-1998	Instructor Seoul National Politech University Seoul, South Korea
2000-2003	Graduate/Project Assistant, College of Education Florida International University Miami, Florida
Present	Doctoral Candidate Florida International University Miami, Florida

PUBLICATIONS AND PRESENTATIONS

Hyun, M. S., (2004). North Korea. In J. J. Slater (Ed.), *Teen life around the world: Teen life in Asia* (pp. 153-168). Westport, CT: Greenwood Press.

Smith, D. H. & Hyun, M. S. (2002, November 21). What's this love affair with the Internet? The use of blended learning in E-Learning. Presented at the 51st National Adult and Continuing Education Conference, St. Louis, MO.