AN ANALYSIS OF SUBGROUP DIFFERENCES ON SELF-ASSESSMENT SCORES OF COLLEGE READINESS SKILLS

by

RONALD LEE YOUNG

A DISSERTATION

Presented to the Department of Educational Methodology, Policy, and Leadership
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Education

June 2014

DISSERTATION APPROVAL PAGE

Student: Ronald Lee Young

Title: An Analysis of Subgroup Differences on Self-Assessment Scores of College Readiness Skills

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Education degree in the Department of Educational Methodology, Policy, and Leadership by:

David Conley	Chair
Gina Biancarosa	Core Member
Keith Hollenbeck	Core Member
Audrey Lucero	Institutional Representative

and

Kimberly Andrews Espy	Vice President for Research and Innovation; Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2014

DISSERTATION ABSTRACT

Ronald Lee Young

Doctor of Education

Department of Educational Methodology, Policy, and Leadership

June 2014

Title: An Analysis of Subgroup Differences on Self-Assessment Scores of College Readiness

College and career readiness is an essential element in the success of students post high school; predicting how successful a high school student will be in a post-secondary setting is the focus of substantial research. Many tools and methods exist for predicting a student's readiness for college and career; unfortunately, gaps persist between readiness rates of different groups of students. The purpose of this study is to examine diagnostic data generated by an innovative survey tool to determine the relationship between high school students in ninth and tenth grades and their self-assessments of five subscales measuring college readiness.

Using extant data collected by the Educational Policy Improvement Center as part of the *CampusReady* tool, this study uses basic descriptive and inferential statistics to look for differences between groups. Results suggest significant differences in the way that students from certain populations (students whose parents have completed differing levels of education, economically disadvantaged students, and Hispanic students) self-evaluate key college readiness skills. Findings from this study will inform K-12 practitioners who plan/develop college and career readiness programs. The impact of

student self-reporting of college and career readiness has implications for future studies that aim to integrate college and career readiness programs.

CURRICULUM VITAE

NAME OF AUTHOR: Ronald Lee Young

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
University of Maryland, College Park
Washington College, Chestertown, Maryland

DEGREES AWARDED:

Doctor of Education, Educational Leadership, 2014, University of Oregon
Master of Arts, Curriculum and Instruction, 2010, University of Maryland
Bachelor of Arts, English and Education, 2006, Washington College

AREAS OF SPECIAL INTEREST:

Career and College Readiness
Educational Leadership Competencies
Issues of Educational Equity

PROFESSIONAL EXPERIENCE:

Principal, The International School, Portland, OR, 2013-Present

Coordinator, Talented and Gifted Education, Sherwood School District, Sherwood, OR, 2011-2013

Teacher, Department of English, Queen Anne's County High School, Centerville, MD, 2006-2011

GRANTS, AWARDS, AND HONORS:

Ken W. Ericksen Memorial Award, University of Oregon, 2012-2013

Magna Cum Laude, Washington College, 2006

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to Professors Conley, Biancarosa, Hollenbeck, and Lucero for their assistance in the preparation of this manuscript. In addition, special thanks to Dr. Terri Ward for her early support of my work. I would be remiss not to mention the members of my cohort, who each provided some motivation, humor, and intellectual challenge throughout the years of the program. I would also like to thank my colleagues in the Sherwood School District and at The International School for listening to my triumphs and tribulations as I have completed this work.

To Mom, Dad, and Corey for their love and support; to Devin for everything always.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Upon taking office in 2008, the Obama administration set a clear college readiness goal that has greatly influenced the trajectory and speed of college readiness entering the national educational discourse (Castro, 2013). In fact, educational scholars and professionals have increasingly sounded the alarm around the lack of adequate preparation of high school graduates for success in postsecondary education and the world of work. The reason for such alarm stems from the important implications of not being ready to successfully complete a college degree, including an inability to effectively access economic, political, and social opportunities (Greene & Forster, 2003; Porter & Polikoff, 2012).

The economic effects of not completing a college degree are wide-ranging. Labor economists project that 62% of U.S. jobs in 2018 (as compared with just 23% in 1973) will require education beyond high school (Rothman, 2012). These jobs require highly sophisticated skills at all levels of industry, skills that can often be learned and refined in a college classroom (ACT, 2004). When people do not complete a postsecondary degree to learn such skills, they are not able to reach full income potential. For example, in 2006, men with a bachelor's degree had an average annual income of $80,942 as compared to an average annual income of $37,356 for a male high school graduate; the income disparity was similar among women, $48,000 vs. $23,000 respectively (Combs et al., 2010). Another example of the economic impact of poor college preparation is that students who enroll in post-secondary programs using federal student aid, but who do not complete a degree within six years, sacrifice, on average, 35% of their annual income to

student loan debt (Sparks & Malkus, 2013). Each of these examples highlights an important reason to successfully complete post-secondary education: not doing so can be economically damaging.

Statement of the Problem

The rates of those in the United States who have completed postsecondary degrees and can avoid these economic and other pitfalls are low. According to Rothman (2012), in 2009, the U.S. college graduation rate was 49%, pushing the U.S. to rank 15th out of 20 major industrialized countries for numbers of people with bachelor's degrees. And the numbers of high school students who graduate and then enroll in college are also low: only 75% of American high school students earn a high school diploma in four years and only 64% of those who graduate attend a 2 or 4-year institution of higher learning (College Board, 2011). Given these statistics, it appears there is a college readiness problem in the United States.

Evidence of inadequate preparation. A potential reason for the smaller numbers of those completing college degrees is the inadequate preparation of students in high school. Rates of enrollment in remedial courses, non-credit bearing classes offered for students who lack the reading, writing, and math skills required for college-level work, are sometimes cited as evidence of inadequate preparation for success in college (Attewell et al., 2006). According to a recent report from the National Center for Education Statistics (Sparks & Malkus, 2013), the aggregate percentage of first-year college students who enroll in at least one remedial course at institutions of higher learning is about 20%; these rates are significantly higher for African American and

Hispanic students (30% and 29%, respectively)[1]. When examining the rates of students enrolling in remedial courses in their first and second years (at degree and non-degree granting programs), the aggregate percentage of students increases to 38% (College Board, 2011). This means that of a class of 100 high school seniors, only 75 will graduate in four years, only 48 will enroll in a two or four-year institution, and only 30 of those students will avoid remedial courses.

In addition to indicating the under-preparedness of college attendees, enrollment in remedial course work has been shown to have negative financial consequences for both the student and the broader community. First, there is an additional financial burden placed on students who must enroll in remedial or "developmental" coursework; these courses often cost the same amount or more per credit and do not count toward a student's degree, thereby increasing the amount of money that a student must pay or must acquire in loan debt (Moore et al., 2010). Furthermore, if the student is receiving federally subsidized loans, the cost of providing these subsidies has been estimated to cost taxpayers around $3 billion annually (Complete College America, 2012). These facts are compelling some policymakers to argue that the prevalence of remedial coursework among college students is evidence that many of these students are not academically prepared to complete college-level work and should not have been admitted to college in the first place (Attewell et al., 2006).

Finally, researchers have pointed to a disturbing relationship between enrollment in remedial coursework and unfinished degrees (Moore et al., 2010). Generally, remedial courses are taken without credit being granted; therefore, a student who enrolls in

[1] These data are for the 2007-2008 academic year, the most recent year available.

multiple remedial courses in the first year of college may not accumulate the same number of credits as his/her peers. According to Adelman (2006), students who completed less than 20 credits by the end of the first calendar year of enrollment were less likely to complete a degree. Attewell et al. (2006) found that taking remedial courses lowered the average chance that a four-year college entrant will graduate by about seven percent. Ultimately, these statistics underscore the implications of not being ready to succeed in postsecondary education.

Defining and measuring college readiness. If completing college is important for success later in one's life, it follows that significant effort might be invested in predicting how college-ready a person may be before he or she enrolls. Though there is wide variation in the components included, college readiness can be defined generally as the degree to which a student's skills and experiences have prepared him/her to successfully meet the essential demands and expectations of a two or four-year postsecondary institution (ACT, 2004; Conley, 2008; Greene & Forster, 2003; Porter & Polikoff, 2012). In order to measure the degree to which students' skills and experiences have prepared them for success in postsecondary classrooms, a variety of predictive approaches have been developed. These predictors can be grouped according to the approach each takes; they include standardized assessment, academic history, and psychosocial factors (Barnes, Slate, & Rojas-LeBouef, 2010; Porter & Polikoff, 2012).

Unfortunately, many of these measures of college readiness are limited in their efficacy. In some cases, these predictors do not measure the full range of what students need in order to be successful in postsecondary education (Conley, 2008). In other cases, students from traditionally underserved subpopulations are under-represented or do not

participate at all in these readiness predictors (Kirst, 2004). In order to address these potential problems, a more comprehensive definition of college readiness, which emphasizes a broader range of skills, and which calls for a different means of assessment is required (Conley, 2008).

Purpose of the Study

This study weighs in on the debate about predicting and measuring college readiness. More specifically, the purpose was to examine the self-assessment of college readiness skills of different groups of students. The study focused on one of four dimensions of college readiness by examining the self-assessment scores of ninth and tenth grade students for five different subscales from that dimension. The specific focus of the analysis was to explore the differences that may exist between the groups of students and their self-assessment of college readiness.

Theoretical Framework

A more comprehensive definition of college readiness makes an attempt to unite the many features of postsecondary education and the ways in which students need to prepare for success. The theoretical framework for this study: the Four Keys to College and Career Readiness developed by Conley (2007), works toward this goal. In this definition, the critical dimensions of readiness are key cognitive strategies, key content knowledge, key learning skills and techniques, and key transition knowledge and skills. This alternative model of college readiness avoids using single data sources (e.g. cut scores) to make judgments about student readiness and allows for more individualized college readiness preparation whenever possible (Conley, 2012). It is also important to note that more recent work on the topic of college readiness (see Porter & Polikoff, 2012)

often uses Conley's theoretical model of college readiness, making it an appropriate choice for this study.

The Key Cognitive Strategies dimension identified in the Four Keys model is of particular interest in this study because of the markedly different approach to college readiness that its inclusion represents. As will be explored in the next chapter's review of the literature, cognitive strategies or the "ways of thinking" (Conley, 2012) are not consistently addressed in the most traditionally accepted predictors of college readiness. For example, ACT, a commonly accepted college readiness test, is designed to measure curriculum taught in high school rather than cognitive processes (Cimetta, D'Agostino, & Levin, 2010). In addition, the shift of focus onto cognitive processes is also reflected in larger curricular reforms currently affecting public K-12 education (Conley, 2011).

An important component of the Four Keys model is the use of a variety of assessment approaches to measure student preparedness; one of these approaches relies on the self-assessment of students on items that measure their preparedness for each of the key dimensions. The *CampusReady* tool (Educational Policy Improvement Center [EPIC], 2011) is an online assessment of college and career readiness used by schools to assess the skills of high school students in the four key areas. *CampusReady* is a unique opportunity because it offers students a chance to self-assess the components of all aspects of college readiness represented by the four dimensions.

Research Questions

In order to explore the potential differences, the study attempted to answer the following research questions using the *CampusReady* self-assessment score data from

ninth and tenth grade students from high schools across the Northwest of the United States. The research questions were

1. What are the associations among the five subscale score variables: communication, interpretation, problem formulation, research, and precision/accuracy?
2. Is there a mean difference in Key Cognitive strategies subscale scores between subgroups of students?

The second research question is further divided into a series of sub-questions that identify the specific subgroups of interest. More specifically, the study has been designed to determine

a) Is there a mean difference in Key Cognitive strategies subscale score between males and females?
b) Is there a mean difference in Key Cognitive strategies subscale scores between students who qualify for free and reduced meals at school (a proxy for economic disadvantage)?
c) Is there a mean difference in Key Cognitive strategies subscale scores between students who speak English as a first language and those who do not?
d) Is there a mean difference in Key Cognitive strategies subscale scores between students who have identified different levels of education of their parents (which could indicate first-generation college attendee status)?

e) Is there a mean difference in Key Cognitive strategies subscale scores between students who identify as African American, Hispanic, White, or other ethnicities?

The study was intended to be descriptive in nature and will draw no causal inferences as to the potential differences in self-assessment data.

CHAPTER II

LITERATURE REVIEW

The following review of the research literature situates this study within the larger body of scholarly work on college and career readiness. The chapter is organized according to the approach to predicting college readiness on which the research focuses. Included with each of the groups of predictors is a brief examination of the problems that each approach presents; chief among these problems are the gaps that exist between subgroups of students. The chapter concludes with an exploration of the college readiness skills of interest, the role of self-assessment in education, and the role of demographics in predicting college readiness.

Research on Predicting Readiness Using Standardized Assessment

There is a long history of using some form of standardized assessment to attempt to predict academic readiness for success in college. For example, one pair of researchers in 1967 studied the correlation of scholastic aptitude scores with college grades (Porter & Polikoff, 2012). Of the many examples of standardized assessment being used today, perhaps the most prevalent in attempting to predict college readiness are the SAT and the ACT, both of which have been used in the college admission decision-making process for over half a century (Moore et al., 2010)[2]. It is important to recognize that neither of these assessments was originally designed to serve as college readiness predictors; nonetheless, they have been and continue to be used in this way (Porter & Polikoff, 2012).

The SAT Reasoning Test. The SAT Reasoning Test (formerly known as the Scholastic Aptitude Test and the Scholastic Assessment Test) has traditionally served as a

[2] While both tests were initially known by full names and their acronyms, now both tests go simply by the more widely used letters and no longer assert they are acronyms.

college admissions test. Influenced by early intelligence tests developed for the military, the SAT was originally designed to measure aptitude, the readiness to learn or work in a new situation (Cimetta, D'Agostino, & Levin, 2010; Young, 2003). Over time, the test has evolved to occupy an important position in the college admissions process because of its repeatedly demonstrated predictive relationship with college grade point average (Lawrence, Rigol, Van Essen, & Jackson, 2003). For example, Kobrin et al. (2008) found that the SAT Reasoning Test (all three sections considered together) had an adjusted multiple correlation of .53 with first-year college grade point average (FYGPA), suggesting that the SAT is a valid predictor of FYGPA. An update to this study completed by Patterson and Mattern (2013) found only a slightly higher adjusted multiple correlation of .56. Both studies also found that the relatively new writing section has the highest correlation with FYGPA among the three sections considered individually (.54)[3]. Until recently, the predictive relationship of the SAT was used to make admission decisions, but was not used as a primary predictor of college or career readiness at the secondary level. As educational policymakers became more interested in finding standardized assessment indicators for college readiness, the SAT was a logical choice.

In 2007, designers of the SAT Reasoning Test developed, for the first time, national college readiness benchmarks using SAT scores. These benchmarks, developed by Kobrin (2007), predict a 65% probability or higher of getting a FYGPA of 2.7 or higher. The study resulted in benchmark scores that could be used for colleges to predict how successful prospective students might be if they enrolled at the institution. The initial work on the creation of benchmarks was extended in 2011 and validated using a larger

[3] In 2014, the SAT made the announcement that it would be making major revisions to the test, including the removal of the required writing section, making it an optional supplement (College Board, 2014).

more representative sample and reflecting changes made to the test since the original design (Wyatt, Kobrin, Wiley, Camara, & Proestler, 2011). Currently, the SAT Reasoning Test is being marketed as a college readiness predictive tool, causing some states, like Texas, to develop college readiness standards and accountability systems that use the SAT Reasoning Test score (Moore et al., 2010).

The ACT. Like the SAT Reasoning Test, the ACT has a long history of educational assessment for college admission. The test, originally designed in 1959, measures academic skills that students obtain throughout their K-12 education and which are necessary for success in college work, with an original goal of urging curricular reform at the high school level (ACT, 2012; Allen, 2013, Porter & Polikoff, 2012). Over time, research has linked ACT scores with college enrollment, success in first-year course work, first-year grade point average, retention, and many other metrics; for example, a researcher recently found a median correlation of .44 between ACT Composite score and first-year college grade point average across 192 four-year institutions (Allen, 2013). Despite the pains that ACT has gone to distinguish itself from the SAT, the two assessments are remarkably similar and students' scores on both tend to be roughly similar as well (Perez, 2002).

ACT, much like the SAT, has also created a college readiness benchmark system using its most popular tests in four core content areas. Beginning as early as 2005, Allen and Sconing modeled grades in typical first-year college courses as a function of ACT test scores. The benchmarks they developed define the minimum scores in English, mathematics, reading, and science that a student must have in order to have at least a 50% chance of achieving a grade of B or higher in credit-bearing college English composition,

college algebra, introductory social science, and biology courses (ACT, 2012). These benchmarks are emphasized as being predictive indicators for *typical* students at *typical* colleges (Allen & Sconing, 2005). In the year 2013, the percentage of ACT-tested high school graduates meeting all four college readiness benchmarks was only 26; the percent of graduates meeting the English benchmark was 64, while mathematics was 44, reading was 44, and science was only 36 (ACT, 2013). These results would seem to indicate at the very best that graduates are unevenly prepared (i.e. more prepared in certain content areas) and at the very worst that graduates are not prepared for college at all.

Problems with standardized assessment predictors. Using standardized assessments to predict college readiness can be problematic for a number of important reasons. First, standardized assessments like the SAT Reasoning Test have been shown to be most effective at predicting early college performance (i.e. FYGPA) and not as effective for subsequent years (Perez, 2002; Sternberg, 2006). Additionally, ambiguities persist about the utility of these assessments in terms of meeting college admission goals. For example, some college admission offices use these assessments to measure aptitude, achievement, or ability, three related, but different constructs (Stemler, 2012).

Another limitation of standardized assessment is the under-representation of some populations of students and their differential performance on the measures. In 2012, White students accounted for 51% of the SAT test taking population, while African American students accounted for only 13% and Hispanic students for 15% (College Board, 2012). Similarly, native English speaking students accounted for 72% of the test takers versus 12% for students whose first language was not English. And finally, students who would be first-generation college attendees were 36% of the test taking

population while the other 64% were students who would not be first generation (College Board, 2012). In addition to under-representation, students from some subgroups also perform differently than their non-subgroup peers. For example, in 2012, only 5% of African American students and 13% of Hispanic students met the readiness benchmarks in all four subject areas of the ACT, compared to 32% of their White peers (ACT, 2012). These disparities point to a significant limitation of this particular predictor of college readiness: the readiness numbers do not represent all populations of students.

Research on Predicting Readiness Using Academic History

Given the potential challenges to using standardized assessments as predictors of readiness explored in the previous section, it is prudent to explore an alternative set of predictors. One possible alternative group of predictors of readiness includes those that fall into the category of academic history: high school grade point average and classes taken in high school.

HSGPA. High school grade point average (HSGPA) or high school class rank is often cited as a predictor of college readiness (Porter & Polikoff, 2012). In fact, Maruyama (2012) posits that in some instances high school grades may be a more desirable predictor than test scores in order to determine readiness. When considered alone (i.e. as a sole predictor), HSGPA had a correlation with FYGPA of between .55 and .58 in a large representative sample (more than 200,000) of standardized assessment test-takers (Patterson & Mattern, 2013; Radunzel & Noble, 2012). Additionally, HSGPA is cited as being an even stronger predictor when paired with some combination of standardized assessment data. Because HSGPA has been studied so often, many states

have included some measure of HSGPA in their college and career readiness accountability systems (Kallison & Stader, 2012).

Classes taken in high school. Another possible academic history predictor of college readiness is the number of specific courses that a student takes in high school. For example, Greene and Forster (2003) argue that students must complete a minimum sequence of four years of English, three years of math, and two years each of natural science, social science, and a foreign language to be considered college ready. ACT (2004) recommends that high school graduates should take a Core Curriculum: four years of English and three years each of math, natural science, and social science. In 2004, ACT found that only 56% of test takers meet these recommendations. Furthermore, ACT, in the same study, found that if students go beyond the Core Curriculum requirements, their likelihood of scoring higher on the college readiness benchmarks also increases. For example, students who completed the Core mathematics courses (typically, Algebra I, Geometry, and Algebra II) and at least two other upper-level mathematics courses (e.g. trigonometry and calculus) scored 6.9 points higher on the mathematics readiness assessment than their peers who took only the Core Curriculum.

The findings of Greene and Forster (2003) and ACT (2004) are further supported by the decades-long research of Adelman, who in 1999 and 2004 studied data indicating that "the academic intensity of the student's high school curriculum still counts more than anything else in pre-collegiate history in providing momentum toward completing a bachelor's degree" (p. xviii). Using an academic intensity index, a 31 level measure of rigorous courses that students could take in high school, Adelman completed a series of transcript analyses using longitudinal data collected by the United States Department of

Education, matching academic intensity scores of some students with other measures of college success (e.g. degree completion time). To date, Adelman's study provides some of the strongest evidence to support the positive relationship between a sufficiently rigorous high school course experience and later college success; the correlation of curriculum intensity and bachelor's degree attainment is .52 (Attewell, et al., 2006; Porter & Polikoff, 2012). In some states, Adelman's study has helped to inform mandatory graduation requirements that highlight the types of rigorous coursework required for college readiness (Moore et al., 2010).

Finally, Kallison & Stader (2012) completed a study that supports the use of academic history as a predictor of college readiness by exploring the utility of "bridge programs," specially designed programs that target students who may need additional academic support to be successful in college. These programs often occur during the summer months for students entering the 11th and 12th grades and differ with the traditional summer school courses. The bridge courses focus entirely on core academic skills needed for college readiness, while traditional summer school courses focus on recovering credit or remediation. And while the number of participating institutions was small, some bridge programs did have a modest impact on student reading and writing achievement.

Problems using academic history predictors. Despite the frequency with which academic history predictors have been used to indicate college readiness, many significant problems persist with this category of predictors. HSGPA remains an inconsistent and sometimes-suspect metric and high-leverage courses are inequitably available to students across the country.

In order to be used as a national predictor of college readiness, HSGPA would need to be calculated using a common metric and have a common meaning across high schools, which is currently not the case (Porter & Polikoff, 2012). Additionally, concerns about the possibility of grade inflation, the phenomena where mean grades increase over time while the quality of the signal that those grades carry degrades, need to be definitively resolved (Pattison, Grodsky, & Muller, 2013). Without some means of standardization, HSGPA continues to be a problematic indicator.

Like HSGPA, using courses taken in high school as a predictor of college readiness can also be problematic. One major problem is that some specific groups of students often attend high schools that do not provide opportunities to learn this advanced content. For example, Adelman (2006) found that the majority of students who took trigonometry earned a bachelor's degree, but only 60% of Hispanic students even attended a high school that offered the course, thereby depriving almost half of even having the opportunity to learn advanced math. Cates & Shaefle (2011) argue that these disparities persist; in other words, Hispanic students are less likely to have academic histories that include rigorous high school courses and students from low socio-economic status backgrounds are less likely to be enrolled in advanced math and science programs. If these populations of students do not have access to these types of courses, then using their course histories as predictors of readiness is not appropriate.

A second problem with using course history as a predictor of college readiness is the inconsistent instruction in these courses. Much like HSGPA, there is little standardization across the country when it comes to the content of high-leverage courses (Venezia & Kirst, 2005). For example, the content, structure, and rigor of a trigonometry

class in one school district may not be the same in a neighboring district. Without clarity and consistency of content, it is difficult to use course titles to predict college readiness.

Research on Predicting Readiness Using Psychosocial Factors

Another possible alternative group of readiness predictors are those that refer to attitudes, beliefs, and attributes (Conley, 2013). Deciding on a terminology for this group of predictors is problematic considering the various terms that seem to be used interchangeably in the literature. For example, these skills have been termed non-academic, non-cognitive, and affective factors (Conley, 2013; Porter & Polikoff, 2012; Robbins et al., 2006). Generally, the research on other factors that may predict college readiness groups the indicators into psychosocial factors (ACT, 2007; Barnes et al., 2010; Kamarraju et al., 2013).

Psychosocial factors, sometimes called "soft skills," are psychological, emotional, and/or social factors that could predict college readiness. For example, Komarraju et al. (2013) found that academic discipline, defined "as the amount of effort students put into school work and the extent to which they see themselves as hardworking" (p. 2), can be predictive of success in college. In the study, Komarraju and colleagues found that academic discipline accounted for nearly 14% of the variance observed in college students' grade point average. In fact, researchers argue that this indicator is just as valid as traditional academic indicators (e.g. HSGPA and ACT scores) (ACT, 2007). Robbins et al. (2006) also found academic discipline to be the most predictive psychosocial factor of college success. In another study, self-efficacy, a student's belief in his or her own ability to succeed, when considered along with cognitive ability and past performance,

may also be a moderate predictor of college success, with correlations ranging from .38 to .50 (Brown et al., 2008).

Problems using psychosocial factors. Like with previous predictors of readiness, it is important to be cautious when considering psychosocial factors as predictors of college readiness. Measuring psychosocial factors can be problematic, although interesting models are available (see Robbins' Student Readiness Inventory). How to contextualize the assessment of psychosocial factors remains an area requiring more research (Robbins et al., 2006). Additionally, research that has examined this category of readiness predictors has not followed students beyond the first or second year of college, so there is limited evidence about the efficacy of predicting long term success using these factors (Cates & Schaefle, 2011).

Research on the College Readiness Skills of Interest

Each of the previous sections has examined groups of readiness predictors that have been used by practitioners and researchers to anticipate a high school student's readiness for success in college. Some of these predictors focus solely on the skills learned in high school, mastery of high school curriculum, or psychosocial skills that could potentially predict readiness. This section of the literature review makes a shift from traditional predictive approaches to a newer approach rooted in an important theoretical model in the field. Based on over 18 years of work on the subject of college readiness, Conley (2007) developed the Four Keys to College and Career Readiness model. Included in this model are four critical dimensions: key cognitive strategies, key

content knowledge, key learning skills and techniques, and key transition knowledge and skills.[4]

Why Key Cognitive strategies? This study focuses on one dimension of the Four Keys model, the Key Cognitive strategies, because of the central importance of this dimension in predicting college readiness. According to Dzubak (2010), students who enter college and struggle are often ill prepared for the "thinking" that occurs in college environments. In fact, some studies have found that college faculty frequently cite unpreparedness for intellectual demands as the core deficiency among first-year college students (Conley, 2007). It seems critical therefore to specifically articulate the skills and strategies that are most illustrative of the "ways of thinking" (Conley, 2012) necessary for success in college. The Key Cognitive Strategies are defined as

> patterns of intellectual behavior that lead to the development of skills and capabilities necessary for college level work. They enable students to learn, understand, retain, use, and apply content from a range of disciplines and are developed within the ways of knowing a particular content area (Conley et al., 2010, p. 25).

It is critical to note that these skills are the underlying cognitive processes that occur when students learn new information; it is how students learn new content. It seems pertinent to focus on this particular component in the Four Keys model because of additional changes in curricula in the public K-12 system. For example, reforms in the states, which have adopted the Common Core standards, have placed a heavier focus on

---

[4] In the initial conceptualization of the framework the areas were: contextual skills and awareness, academic behaviors, key content knowledge, and key cognitive strategies. Older research will have these previously used terms.

cognitive processes in curricula than previous iterations of content standards (Conley, 2011). In addition, the focus on these skills has been supported by research that specifically identifies them as lacking in today's college populations and work force. For example, Knox (2006) claims that only 13% of American adults are proficient in informational skills (e.g. finding, understanding, and using information).

The Key Cognitive Strategies section of the Four Keys model includes the specific skills of problem formulation, research, interpretation, communication, precision and accuracy (Conley, 2012). These skills reflect the "ways of thinking" that students may be called on to do when in college. What follows is an exploration of these skills and connection to the broader research literature.

Problem formulation. The first of the key college readiness skills is problem formulation. The ability to formulate and solve problems has been identified as a critical skill for success in academics and life in the 21st Century (Saiz & Rivas, 2011; Wagner, 2010). Dutch researcher van Merrienboer (2013) explicates three definitions of problem solving: (a) *weak methods* that are domain-general, high-cost, and often ineffective; (b) *strong methods* that are domain-specific, algorithmic, and effective in well-structured problems; and (c) *knowledge-based methods* that are appropriate for ill-structured problems that require multiple interrelated skills to solve. In his research, van Merrienboer argues that modern educational settings need to focus on the third definition, *knowledge-based methods*, which provide the most cognitively rigorous experience for students and which may better resemble the types of problems that students will encounter in their futures. This position is well-supported by Shook Cheong (2005) who posits that contemporary problem solving requires a complicated set of interconnected

skills that require a person to be analytical, creative, and practical simultaneously. Athanasou (2012) further echoes van Merrienboer by defining problem solving as the "goal-directed thinking and action in situations for which no routine solution is available" (p. XX). Of note in this definition is the mention of the novelty of problem situations; people need the skills to clearly recognize the context within which the problem presents itself and the specific demands that any solution would need to include in order to be successful in that situation.

The Problem Formulation component of the Four Keys model of College and Career Readiness reflects the key components of the definitions of problem solving explored above. The model defines problem solving as problem formulation: "when students understand what the problem is asking and the concepts that are being addressed", when they "hypothesize about what might be a plausible answer or preliminary thesis… [and] finally strategize about how to address the task" (Conley et al., 2010, p. 25).

Communication, interpretation, research, and precision/accuracy skills. After problem formulation, the Four Keys model identifies four additional subcomponents to the Key Cognitive Strategies dimension, which include research, interpretation, communication, and precision/accuracy (Conley, 2012). Research is defined as

> the ability to identify the information needed to solve a problem, the collection of sufficient relevant data or sources to answer the question, and the evaluation of sources or data collected to determine their validity, credibility, and relevance, noting any potential sources of error or bias. [Interpretation is defined] as the ability to integrate data or evidence, organizing it in ways that further analysis,

> describing patterns or details, and synthesizing information to offer a justification for conclusions. Communication refers to the ability of students to discern an appropriate way to share information and to construct work products that meet the demands of a specific task (Conley et al., 2010, p. 26).

The final aspect, precision/accuracy, is defined as a student's ability to be "appropriately precise and accurate at all stages of the process by determining and using language, terms, expressions, rules, terminology, and conventions appropriate to the subject area and problem" (Lombardi et al., 2013, p. 165). These related metacognitive skills can be grouped under a modern version of the concept of information literacy.

The traditional understanding of information literacy was defined within the context of library science and often referred to an understanding of call numbers/other cataloging systems and the location/interpretation of various resources throughout a library (O'Connor et al., 2002). However, more modern understandings of the term define information literacy as a broader ability to find and locate information. For example, Marcum (2004) explains that the ability to access, manage, and utilize information to make decisions and solve problems is information literacy for the new age. Marcum's definition is echoed by the Partnership for 21st Century Learning, a think tank that defines information literacy as the ability to

> access information efficiently and effectively; evaluate information critically and competently; use information accurately and creatively for the issue or problem at hand; manage the flow of information from a wide variety of sources; and apply a fundamental understanding of the ethical/legal issues surrounding the access and use of information. (2011).

These definitions of information literacy align with the subcomponents (research, interpretation, communication, and precision/accuracy) from the Key Cognitive Strategies category of the Four Keys Model and have been validated as critical skills for college and career readiness (Knox, 2006; Wagner, 2010).

The Role of Self-Assessment

Current predictors of college readiness that fall in the traditional categories of standardized assessment, academic history, and psychosocial factors do not directly address some key readiness skills. Yet the comprehensive measurement of skills and/or student performance is critical to developing policies to improve teaching and learning at the secondary level (Porter & Polikoff, 2012). In addition to missing critical college readiness skills, current predictors also miss an opportunity to include student voice and allow students the opportunity to examine themselves and their progress toward becoming college ready (Heritage, 2009). An instrument which includes some degree of self-assessment and which focuses on specific skills may offer a unique opportunity to examine student readiness.

Self-assessment has been widely used in educational literature to refer to various concepts and/or academic processes. Within the broad umbrella of self-assessment there are three distinct activities: *self-testing,* checking performance against provided test items; *self-rating,* appraising the present state of one's knowledge, skills, or achievement; and *reflective questioning*, reflecting on learning through open-ended questions (Boud & Falchikov, 1989; Ibabe & Jauregizar, 2010). These activities can be used to gather information that sometimes cannot be gained by other methods (Smyth & Terry, 2007). In addition, self-assessment methods can benefit learners by allowing them to become

more autonomous agents in their education, taking responsibility for gaining and improving on their knowledge and skills (Dunning, Heath, & Suls, 2004). Finally, concerns about whether or not self-assessment leads students to inflate the appraisals of their own work can be potentially ameliorated by framing the self-assessment task appropriately, emphasizing the relatively low stakes (Andrade & Du, 2007).

While evidence has been found to support the use of self-assessment in both formative and summative assessment tasks (Heritage, 2009), empirical evidence is limited. For example, Chang, Liang, and Chen (2012) examined the use of self-assessment at the high school level and found evidence of reliability and validity in that context, especially for assessment tasks that are ongoing. Earlier work also indicated that students with more educational experience (e.g. students in high school) could provide more accurate self-assessment (Dunning, Heath, & Suls, 2004). The role of self-assessment is clearly important in the process of learning, well beyond that of simply assigning grades or marking knowledge (Kirby & Downs, 2007).

Self-assessment as a method of exploring and/or predicting college readiness is not necessarily new. Self-assessment tools have been more commonplace in measuring some of the skills for college readiness that fall into the psychosocial category (Robbins et al., 2006; Lombardi, Seburn, & Conley, 2011). EPIC has produced a specially designed self-assessment tool that explores the readiness skills represented in the Four Keys of College and Career Readiness model. The *CampusReady* is an instrument that includes a series of surveys for teachers, administrators, counselors, and students, and relies on over 1100 items. *CampusReady* provides analytical tools for leaders and policy makers to examine the impact of school policies, teacher/counselor practices, and the

self-assessment of students (Conley et al., 2010). Recent research has validated the use of the *CampusReady* instrument to assess the Key Cognitive strategies, providing evidence that there is preferable reliability and promising validity with alpha levels ranging from .88 to .93 (Lombardi, Conley, Seburn, & Downs, 2013).

Subgroup Differences

Like many other areas of education, there is collective professional and scholarly concern about the college readiness of particular subpopulations of students. The following section of the literature review explores the subgroups of students that will be examined in this study and synthesizes some of the current research on the college readiness, enrollment/attendance, and success for each.

Gender. One area in which the college readiness of groups is explored is gender differences. Males and females have been shown to differ in their readiness for college and their success while in college. In terms of readiness, a statewide study of students found that females were more prepared than males for the level of text complexity found in college (Wilkins et al., 2012). However, other studies have found female students are less likely to be ready for college-level math than their male peers and are less likely to access higher-level math courses in high school than males (Long, Iatarola, & Conger, 2009; Combs et al., 2010). This phenomenon is also evident in SAT scores on the quantitative section, where males have generally had higher scores than females (Nankervis, 2011).

Once in college, differences persist. In 2007-08, 25% of females reported enrolling in remedial coursework, while only 22% of males did the same (Snyder, Dillow, & Hoffman, 2009). Ganzert (2012) found that in one statewide sample, males

were more likely to take advantage of dual-enrollment programs while in high school, but females were more likely to have a higher FYGPA. Additionally, and perhaps most importantly, since 1991, more women than men have enrolled in and completed four-year degree programs at postsecondary institutions, a phenomenon that is present among all racial and ethnic categories (Population Reference Bureau, 2013).

Socioeconomic status. Perhaps one of the most often studied group differences in college readiness has been among socioeconomic groups. A student's status as a member of the low-income socioeconomic group can have implications in terms of high school completion, college enrollment, and college success.

In a meta-analysis of 58 studies concerning secondary educational achievement and socioeconomic status, Sirin (2005) found an effect size of $d = .66$ between achievement and free and reduced meals status (FARMS), indicating that FARMS can have real effects on student academic achievement (Hattie, 2009). The differences between socioeconomic groups continue in college. Only 52% of low-income students immediately enrolled in college after graduating high school in 2004 compared to almost 80% of high-income students (Roderick, Nagaoka, Coca, 2004). Even when they have the necessary qualifications, these students are less likely to attend college (students from low-SES backgrounds applied to college at a rate 17% lower than the national average for similarly qualified students) and they are less likely to be academically prepared for college-level work (Cates & Schaefle, 2011). One potential reason for the lower enrollment rates in postsecondary institutions could be because these students possess less knowledge about the general process of applying and attending college than their higher-income peers (Roderick et al., 2004). Once in postsecondary settings, these

students are disproportionately represented in community colleges and in remedial education courses and are less likely to complete a four-year degree (Castro, 2013).

English language learner status. Students for whom English is a second language and who are enrolled in public schools in the United States are one of the fastest growing groups of students in the country, with a total population in 2010-11 of approximately 10% of all public school students (US Dept. of Education, 2013). These students perform well below their native-English speaking peers on measures of achievement in K-12 schools; for example, NAEP reading scores at the 8th grade were 36 points lower for ELLs (US Dept. of Education, 2013). Furthermore, schools with the highest percentage of English language learners also had the fewest number of AP classes (Yun & Moreno, 2013). Additionally, ELLs have a very low enrollment percentage in postsecondary institutions: 13.5% compared to 37% of native English speaking peers in 1999 (Kanno & Cromley, 2013). Yet, the most striking fact about college readiness rates for English language learners is the fact that there are no national-level statistics on ELL's college access and attainment in the United States (Kanno & Cromley, 2013).

Level of Educational Attainment of Parents. Unlike ELL status, an abundance of information exists in the research literature about the college readiness of students whose parents have not attended college. Whether or not a student will be a first-generation college attendee can influence key areas of college readiness including pre-collegiate traits, college experiences (both academic and non-academic), and educational outcomes (including degree completion) (Hahs-Vaughn, 2004).

One important pre-collegiate trait that is affected by parental educational attainment is knowledge about college. A family's awareness about college, the amount

and consistency of information high school students and their families receive about preparing to take college-level courses, can influence their level of preparation (Venezia & Kirst, 2005). Additionally, parents' educational levels may mean that students do not receive information about college-track curricula at home (Cates & Schaefle, 2012). This lack of knowledge about college and its importance can also cause an aspirations gap because students whose parents have attained no more than a high school diploma are the least likely to aspire to a bachelor's degree and least likely to be college qualified, even when controlling for other factors like academic achievement and socioeconomic status (Hahs-Vaughn, 2004).

If they make it to college, first-generation college attendees are more likely to need remediation; in 2007-08, the highest percentage of remedial course-takers were students whose parents had only completed high school (Snyder, Dillow, & Hoffman, 2009). First generation students work more hours than their non-first generation peers, and work responsibilities have a stronger negative effect on their college experiences including critical thinking, internal locus of control in relation to college success, and preference for cognitive tasks that engage higher order thinking skills (Pascarella et al., 2004 qtd. in Hahs-Vaughn, 2004). All of these factors combine, making success completion of a bachelor's degree less likely for first-generation college students, a reduction of roughly 21% (Adelman, 2006).

Race/Ethnicity. Like the other subgroups of students explored so far, race/ethnicity has an impact on college readiness, enrollment, and success. According to estimates, in 2004, only 58% of Hispanic students enrolled in college immediately after high school, compared to 59% of African American students and 69% of White students

(Roderick et al., 2004). Furthermore, the differences between groups on the traditional predictors of college readiness are large. For example, in 2005, the cumulative high school GPA of White students was 3.05, while Hispanic students' was 2.82 and African Americans was 2.69 (Roderick et al., 2004). Additionally, Hispanic students, in particular, are less likely to have access to high-leverage academic courses (e.g. advanced mathematics or advanced science) (Cates & Shaefle, 2012; Roderick et al., 2004). Ethnic minority students are also less likely to possess the same level of knowledge of the college application process, the financial aid system, and the range of choices within the postsecondary system than their White peers (Roderick et al., 2004; Venezia & Kirst, 2005). These facts each lead to an overall lower college-readiness rate for students in ethnic minority groups; 16% of Hispanic students can be considered college ready according to Greene & Forster (2003), while 20% of African American students are ready and 37% of White students.

Further amplifying the effects of race/ethnicity on college readiness is the fact that often times the same students are within an ethnic minority and are considered low-income, thereby concentrating their educational disadvantage (Yun & Moreno, 2013). Schools with the highest percentage of African American and Hispanic student populations also had the fewest number of experienced teachers and AP classes (Yun & Moreno, 2013). The quality of educational experiences for Hispanic students at the secondary level, in particular, is less than that of their White peers; perhaps leading to higher rates of remediation among Hispanic students (28%) versus their White peers (20%) (Cates & Schaefle, 2012; IES, 2013).

Summary

To summarize, the research on college readiness has developed over time and has turned to a variety of predictors. Standardized assessments have been used to alternatively measure aptitude, achievement, and curricula needed for college. Academic history predictors have been used to represent student academic experiences and performance. Psychosocial factors have been used to predict success based on constructs like academic discipline, persistence, motivation, and the like. Each of these traditional predictors has strengths and shortcomings. But one area that has been used less often is that of self-assessment, which allows students the voice with which to express their own appraisals of their readiness. Focusing on one dimension of a comprehensive model of college readiness, the Key Cognitive strategies from Conley's Four Keys model, this study will examine self-assessment data, but through the important lens of educational equity, searching carefully for the same differences among subgroups that have been shown in other predictors.

CHAPTER III

METHODOLOGY

The previous chapter reviews the relevant research literature on predicting college readiness and makes the persistent argument that a comprehensive definition of college readiness, along with alternative means of assessment, needs to be adopted by scholars, policymakers, and professionals. This chapter explores the methodological approach used to conduct this study, with special emphasis on the analysis of data. The purpose of this study is to use basic descriptive and inferential statistics to investigate the potential differences between subscale scores for groups of ninth and tenth grade students. More specifically, the study addressed the following main research questions:

1. What are the associations among the five subscale score variables: communication, interpretation, problem formulation, research, and precision/accuracy?
2. Is there a mean difference in Key Cognitive strategies subscale scores between subgroups of students?

The second research question is further divided into a series of sub-questions that identify the specific subgroups of interest. More specifically, the study has been designed to determine

a) Is there a mean difference in Key Cognitive strategies subscale score between males and females?
b) Is there a mean difference in Key Cognitive strategies subscale scores between students who qualify for free and reduced meals at school (a proxy for economic disadvantage)?

c) Is there a mean difference in Key Cognitive strategies subscale scores between students who speak English as a first language and those who do not?

d) Is there a mean difference in Key Cognitive strategies subscale scores between students who have identified different levels of education of their parents (which could indicate first-generation college attendee status)?

e) Is there a mean difference in Key Cognitive strategies subscale scores between students who identify as African American, Hispanic, White, or other ethnicities?

Data have been collected from the *CampusReady* diagnostic tool, a comprehensive self-assessment tool, created and managed by the Education Policy Improvement Center. This chapter addresses the methodological approach to answering the research questions listed and includes a description of the design, population, sample, instruments, and analyses to be conducted.

Study Design

The design of this study is a secondary analysis of extant data using basic inferential and descriptive statistics. The independent variables in this study are the subgroups in which students have identified themselves. Included in the data are demographic items that can discern a student's gender, race/ethnicity, free and reduced meals status (a proxy for economic disadvantage), highest education level attained by the student's parents, and English language learner status. The dependent variables are the scores of each of five subscales in the Key Cognitive strategies dimension of the

CampusReady instrument. The study was conducted in phases to explore the nature of the differences that exist between groups.

In the first phase of the study, a simple correlation was completed to be certain that subscales were related to one another. Though the subscales are conceptually very close, all five are aspects of the same dimension of college readiness; it was critical to establish statistical association for later statistical tests.

The second phase of the study examined differences among subscales for various groups of students. Because the second research question (and series of sub-questions) essentially is interested in predicting an outcome based on membership in a group, basic statistical procedures that compare means were selected. This phase included several independent-samples *t*-tests and a series of ANOVA tests in order to consider differences in the dependent variables of different subgroups. The *t*-tests and ANOVAs are the appropriate statistical tests because both are robust and allow practitioners to quickly discern the presence of statistical difference. An effect size measure was calculated for each *t*-test and ANOVA in order to determine the practical significance of the findings (i.e. to determine the magnitude of difference).

There are potential disadvantages to using the ANOVA during the second phase of the study. For example, unequal group sizes can have an important effect on creating biased results. ANOVA is most accurate when there are at least 20 degrees of freedom and the smallest response category contains at least 20% of all responses; although if the smallest category has fewer than 20%, the ANOVA can be accurate with 40 or more degrees of freedom (Field, 2013).

Population and Sample

The sample for this study comes from extant data collected by the Education Policy Improvement Center, through its *CampusReady* diagnostic tool. Respondents in the sample come from 27 high schools spread across 15 different school districts across the Northwest portion of the United States. All responses were collected between 2009-2011. The data include student self-assessment scores for items assessing all aspects of the four dimensions of college readiness as well as demographic data for each student.

The self-assessment scores and demographic information was collected from school administrators, counselors, middle school teachers, high school teachers, middle school students, and high school students. EPIC provided schools with directions on how to identify potentially random samples from their student bodies to complete the diagnostic. Once random samples were identified, participants were compelled to participate. For the purposes of this study, only 9th and 10th grade students' responses are considered.

Of the total number of respondents, 3154 were students in the ninth and tenth grades. Table 1 presents the demographic information on the total sample, including the number of participants from each of the subgroups of interest. The majority of the group was White and native English speaking. More students indicated that they qualified for free and/or reduced school meals (either breakfast or lunch) than those who do not qualify and almost ¾ of the students would be first-generation college attendees.

In order to meet requirements for use in this study, respondents needed to have answered at least 50% of the items measuring each of the five aspect subscale scores. This rule greatly reduced the sample size for each subscale and within each subgroup. The adjusted sample sizes for each subgroup can be found in Chapter IV of the study, under descriptive statistics. In addition, if students did not answer enough items for any of the subscales, they were not included in the study. The data also present problems because of unequal group sizes, although this particular issue was considered and appropriate statistical procedures were used that account for this problem (see analyses section of this chapter).

Table 1

*Demographics of the Total Sample of Ninth and Tenth Grade Students (*N = *3154)*

Characteristic	*n*	%
Gender		
Male	1567	49.7
Female	1585	50.3
Ethnicity		
White	1366	43.3
African American	366	11.6
Hispanic	817	25.9
Other	605	19.2
English as First Language		
Yes	710	22.5
No	2411	76.4
DKNA	30	1.0
First-Generation College Attendee		
Yes	2204	69.9
No	633	20.1
DKNA	315	10.0
Free & Reduced Meals Qualified		
Yes	1620	51.4
No	1432	45.4
DKNA	99	3.1

Survey Instrument

The *CampusReady* is a tool that has been developed by the Educational Policy Improvement Center, an organization that provides research and tools to empower states, districts, schools, and teachers to prepare students for success beyond high school (EPIC, 2013). The *CampusReady* "improves college and career readiness for students through a self-diagnostic online tool that provides data, reports, recommendations, and links to resources that help schools" (EPIC, 2013). A series of surveys, included in the tool, measure the Four Keys to College and Career Readiness through a web-based diagnostic. Respondents to the surveys include students, counselors, and administrators in middle (grades 6-8) and high schools (grades 9-12).

The Four Keys to College and Career Readiness, alternatively described in some literature as the four-dimensions of college and career readiness, include: key learning skills and techniques, key transition knowledge and skills, key content knowledge, and key cognitive strategies (Conley et al., 2010). The development of this concept was initiated by Conley et al. (2010) in a qualitative study that explored the college readiness practices of 38 high schools across the country matched with an extensive review of the research literature on college and career readiness. Through that research, Conley and colleagues were able to create an earlier version of the *CampusReady* tool then known as the College Career Ready School Diagnostic (CCRSD) tool.

This study is concerned with one dimension of the Four Keys model: the Key Cognitive strategies. Students assess themselves on this dimension in five subscales: communication, interpretation, problem formulation, research, and precision/accuracy. Each subscale score is average of multiple individual items on which students rate

themselves. All items on the instrument use the direction, “please indicate how much each item describes you”, using a Likert scale that ranges from “0” (don’t know/not applicable) to “5” (very much like me). Table 2 presents the number of items that comprise each subscale along with the specific text of each item.

Because this study is a secondary analysis, it was not possible to be present for the original data collection. In order to use the data collected in a meaningful way, the study includes the following assumptions about the participants and instrumentation:

1. Each of the participant schools followed the directions for administration and sampling provided by EPIC.
2. Responses provided by the participants provide accurate, reflective answers to each item.
3. Participants had adequate time and understanding to complete each item of the instrument.

Analysis

Analysis of the data occurred in multiple stages: an exploratory data analysis, correlation analyses between the subscales, and analyses of variance for the subgroups. SPSS 22.0 was the statistical package used for all procedures. Each of these separate stages is explored in more detail in the following section.

The data have been entered, coded, and checked for accuracy by the researchers at the Educational Policy Improvement Center, but for the purpose of this study, the larger sample was divided into a smaller sample of just ninth and tenth grade students.

Table 2

Items Included in the Five Subscales

Subscales of Key Cognitive Strategies (# of Items) Item Text
Communication (12)
27. When studying or approaching a task, I outline material or the task to help organize my thoughts and approach.
28. When writing a paper or making a presentation, I only make points that I can support with evidence.
29. I use logic to build my evidence and defend my point of view.
30. I go through several steps to organize my thinking before I begin writing a paper.
31. I draw upon tools such as outlines or pro and con lists that I have already constructed when I start to write a paper or prepare a presentation.
32. I can construct logical arguments to explain issues or answer questions.
33. I can accept critiques and challenges to claims I've made.
34. I am able to address critiques and challenges to my claims.
35. I make presentations in class.
36. I can provide constructive criticism of others' work.
37. I know how to present my ideas out loud in ways that others can follow.
38. I can present my ideas in writing in ways that others can follow.
Interpretation (11)
19. When an idea is presented in class, I think about the evidence upon which it is based.
20. When I listen to ideas, theories, or concepts discussed in class, I try to decide on my own if there is enough evidence supporting them.
21. I think about the strengths and weaknesses of different versions of issues and events.
22. I try to make connections between what I am learning now and what I have learned in the past.
23. I try to make connections between what I learn in one class and what I learn in other classes.
24. I try to make connections between what I am learning in class and my own experiences.
25. I use evidence to evaluate the strengths and weaknesses of the conclusions I reach.
26. After I complete a large or challenging assignment, I reflect on the final product and identify ways to do similar work better next time.
60. I select different strategies to analyze information, issues, or events based on the nature of the material I am analyzing.
61. Whenever I hear about an assertion or conclusion in class, I think about possible alternative reasons for the assertion or conclusion.
62. When reading for class or conducting research, I look for patterns in information to help make my point.
63. I make charts, tables, or diagrams to help me analyze material for assignments.

Problem Formulation (15)

S53. I know how to draw on my past experience when I try to solve a problem in class.
H53. I know how to identify the purpose of a research problem.
54. I understand an assignment before beginning to work on it.
55. I understand how parts of a research problem interact with each other.
56. I understand how my approach to formulating a research question might change depending on who will see the final results.
57. I develop research questions by identifying what information is available and what information is needed but missing.
58. I know how to generate multiple hypotheses and figure out the best one.
59. I enjoy solving problems that have more than one "right" answer.
01. When faced with a difficult task, I make a plan or strategy to solve it.
02. I know and can use multiple strategies for solving difficult tasks.
03. I think about different ways to solve a complex task and pick the best one.
04. When faced with a complex task, I break it down into smaller pieces.
05. When solving a difficult task, I know when to try a different method if the first methd is not working well.
06. Sometimes I reflect on how I solved a problem, what worked, what didn't, and how I can do better next time.
07. I reflect on my past experiences when thinking about how to solve a difficult task.

Research (11)

13. When I conduct research, I collect information from a variety of sources (books, articles, online, etc.).
14. Before studying new material, I skim to see how it is organized.
15. When I conduct research, I can tell the difference between fact and opinion when I select resources to cite as supporting evidence.
16. When I conduct research, I think about whether my sources can be trusted as credible.
17. When I conduct research, I use information that applies to what I am studying.
18. I am good at coming up with a research plan.
08. Before I solve a difficult task, I think about what information I need to help me find a solution.
09. Before I solve a difficult task, I think about how much research I need to do to find a solution.
10. I look at the information necessary to answer a question or solve a difficult task and identify where I can find it and what I may need to do to collect it.
11. I know what information and resources are available online and how to use them for research.
12. I know what information and resources are available in a library and how to use them for research.

Precision & Accuracy (14)

46. I allow enough time to do a final check of my work before turning it in.
47. I check my work for careless errors before turning in assignments and tasks.
48. I check for spelling and grammar errors before turning in work.
49. I check to make sure I correctly followed all instructions before turning in work.
50. I check to make sure my ideas and conclusions are supported and logical before turning in work.
51. It is important to me to be precise in my schoolwork.
52. It is important to me to be accurate in my schoolwork.
39. I read test questions and instructions carefully before beginning to answer questions.
40. I look for, read, and follow instructions describing assignment deadlines, length, formatting, and purpose.
41. When taking a test, I monitor the time so I am able to complete all items on the test.
42. I listen to feedback about my work and incorporate that feedback into the final versions of assignments.
43. I complete multiple drafts of assignments and make improvements between versions.
44. When different subjects require different writing styles, I make sure the writing style I use is appropriate.
45. In my assignments, I use the vocabulary, notation, or symbols appropriate to the subject area or course.

The sample has been limited to students in the ninth and tenth grades for two reasons: (a) in order to eliminate potential sources of difference in the data and (b) because of the research supporting the use of self-assessment as a formative tool.

In order to minimize the effect of differences caused by grade level, the sample was limited to ninth and tenth grade students. For example, some students at different grade levels may assess their own college readiness skills differently. Because these differences may be attributed to age, maturity level, and/or the number of years of high school study, this potential difference has been minimized. Additionally, research has supported the use of self-assessment as a formative tool (Heritage, 2009); as it would be

more possible to intervene with ninth and tenth graders than twelfth graders, the younger high school grades were selected.

Once the sample was limited to the grades of interest, additional manipulations to the variables were made. First, the Mean function of SPSS was used to find the average subscale score for each of the five Key Cognitive strategies for each respondent. When calculating the mean subscale scores, the minimum number of variables that must have valid (non-missing) values was half the number of total items for each strategy (e.g., for the communication subscale, respondents had to have answered at least 6 of the 12 items measuring that strategy to be included in the study). Once the mean subscale scores were calculated, descriptive statistics were run to check for errors and normality of each new variable. Listwise deletion was used to eliminate respondents who had selected the "don't know/not applicable" choice for demographic data and/or for respondents who did not meet the decision rule described above for the mean subscale scores. The purpose for these deletions was to be certain that the same students were being compared across all five subscales.

In addition to the calculating the subscale scores, two other variables were manipulated before analyses were conducted: the parent's education and race/ethnicity variables. First, the original mother's education and father's education variables were recoded. The original variables included six response options (after removing the "don't know/not applicable" option). The three response options that indicate a high school graduate or less were collapsed together (these included "8th grade or less", "Some high school", and "High school graduate"). Similarly, the two options that indicate some college were combined ("Some college" and "2 years of college"). Finally, the two

options indicating a four-year degree or higher were combined ("Four-year degree" and "Graduate degree"). When these manipulations were completed, the two revised mother's education and father's education variables were combined to create a new variable: parents education. This new variable was defined such that if a student had at least one parent with some college, the "Some college" option was chosen for him/her and if the student had at least one parent with the "Four years or more" option, then that category was chosen. The purpose of these manipulations was two-fold: (a) to increase the sample size of each of the groups, thereby increasing the statistical power and (b) to match the response categories with other research in the field (see IES, 2013).

The second of these additional variables to be recoded was the race/ethnicity variable. In this case, much simpler changes were made. The original set of response options included seven different choices for the race/ethnicity variable. In order to increase the sample sizes for statistical analyses, students who responded "Asian/Pacific Islander", "American Indian/Alaskan Native", or "Multiple Categories/Mixed Race" were recoded into a new variable titled "Other".

Once the variables had been cleaned, calculated, and checked for errors, the exploratory data analysis began. During this stage, the minimum and maximum values for each variable were compared with the allowable ranges of values in the codebook. Basic descriptive statistics were calculated, including measures of central tendency, frequencies, skewness, and kurtosis values.

An important part of the exploratory analysis phase was the examination of critical statistical assumptions: (a) normality; (b) homoscedasticity/homogeneity of variance; and (c) independence. Not only is the testing for each of these assumptions

necessary for the later statistical analyses, but it also decreases the likelihood that additional bias affected the results.

In order to answer the first research question about whether or not there is an association between the dependent variables (subscale scores), a correlation analysis is the appropriate statistical test. Though there is a coherent logic to understanding that the subscales are conceptually related, the correlation analysis allows one to understand if the subscales are statistically related. Assuming that the appropriate statistical assumptions were met (most importantly normality), the Pearson *r* was calculated using SPSS procedures to find the associations between subscale scores.

To answer the second research question about whether differences exist within the data, independent samples *t*-tests and analysis of variance (ANOVA) were the primary method of analysis. The first step in completing this stage of the data analysis is to test whether the appropriate assumptions were met in order to complete these tests. First, observations have to be independent with respect to self-identification in each of the groups represented in that variable. For example, students could either identify as an English language learner or not. Second, the variances on the dependent variable (average Likert[5] score for each item within categories) have to be equal across the groups. Levene's test for homogeneity of variance was used to check for this second assumption. In the one instance when this assumption is violated, the alternative Welch's test is used. Third, the dependent variable has to be normally distributed.

After determining if the data meet the assumptions necessary to proceed in the testing, the actual analyses were conducted to answer the questions pertaining to the

[5] A more in-depth discussion of my treatment of the scale of the Likert variable will follow in the discussion chapter.

existence of a difference between groups. When significant differences are found in the results of the independent samples *t*-tests, Cohen's *d* was used as the effect size measure. When the one-way ANOVAs were completed, and the omnibus *F* statistic was found to be significant, an appropriate post hoc test was conducted to determine where the exact significant differences lay (i.e. among which subgroups). Generally, researchers tend to use Bonferroni's or Tukey's tests as the more commonly used post hoc tests following an ANOVA; however, some research has shown that both perform badly when groups sizes are unequal (Field, 2013). Because the sample sizes are unequal in this study, the Hochberg GT2 test is most often used as the post hoc. Given the sometimes-large differences between group sizes, this post hoc test represents the best compromise between statistical power and control over the error rate (Field, 2013; Toothaker, 1993). After finding significant differences in the post hoc test, the effect size of differences was calculated using Cohen's *d.* When it was not possible to assume equal variance (i.e. because the homogeneity of variance assumption was violated), the Games-Howell post hoc test was used and similarly an effect size was calculated.

Summary

In summary, the purpose of this study was to investigate the differences between subgroups self-assessment of the five aspects of the Key Cognitive strategies dimension of college readiness. The study focused on the subscale scores for each of the aspects of students in the 9^{th} and 10^{th} grades and utilized basic and inferential statistical procedures to uncover the relationships between subscales and the differences that exist within subgroups of students. The results of this study provide potentially important information

to practitioners about the possible use of a comprehensive system for self-assessment of college readiness for students in public high schools in the United States.

CHAPTER IV

RESULTS

This study explored the differences between the subscale scores of ninth and tenth grade students on a self-assessment of college readiness. The sample included students from all over the Northwest of the United States who participated in the *CampusReady* self-assessment program. This chapter begins by reporting whether the statistical assumptions identified in Chapter III were met, then reports descriptive statistics, and finally reports findings related to each of the research questions in sections: correlation and results to the independent-samples t-tests and the series of ANOVAs.

Statistical Assumptions and Requirements

For every statistical model, there are particular assumptions that must be met in order for inferences made from the tests to be considered valid. This study considered three specific assumptions on which most statistical tests are based. The three assumptions are (a) normality, (b) homoscedasticity/homogeneity of variance, and (c) independence.

The first statistical assumption is normality. As Field (2013) notes, the normality assumption can have slightly different meanings depending on the context. For example, considering that this study concerns differences in means, the data do not need to be normally distributed, but the sampling distribution of means does (p. 169). In this case, all dependent variables were approximately normal, when considered as a whole (without taking into account the categorical independent variables). Visual inspection of both histograms and P-P plots indicated an approximately normal distribution for each subscale. In addition, numerical representations for skewness and kurtosis for each of the

dependent variables were also less than 1, indicating an approximately normal distribution. Because the majority of the research questions will examine groups, normality was checked within each subgroup as well. The subgroups of interest are gender, race/ethnicity, economic status (as defined by whether or not a student qualifies for free and reduced meals at school), college attendee status (defined by the highest level of education completed by parents), and English language learner status (defined by whether or not a student speaks English as a first language). Tables 3-7 provide summaries of the descriptive statistics for each subscale and display the skewness and kurtosis statistics. Based on this information, all data were approximately normal.

The second statistical assumption is homescedasticity/homogeneity of variance, which essentially means that the data from the samples come from populations with the same variance (Field, 2013, p. 174). By confirming that there is equality of variance, the estimates of parameters within the chosen statistical model will be optimal. In order to test this assumption, Levene's test for equality of variance was conducted before each ANOVA. When the assumption was not met, an alternative procedure was used to compare means.

The final statistical assumption is independence, which means that the errors in the model are not related to one another. Critical to meeting this assumption is ensuring that respondents did not identify themselves as belong to more than one of the subgroups, thereby making the mean subscale scores relate because they were derived from the responses of the same person. Because the design of the *CampusReady* instrument does not allow students to pick multiple answers within the demographic category (e.g. a

student cannot choose both male and female on the gender question), it is fair to assume that all responses were independent.

Descriptive Statistics

The procedures outlined in Chapter III were followed for this investigation. The Educational Policy Improvement Center (EPIC) provided the data for 9th and 10th grade students from several school districts collected between 2009 and 2011. The total sample size for the study was 888 students. For each subscale the minimum score was .00 and the maximum score was 5.00. Table 3 reports the descriptive statistics for the Communication subscale.

Table 3

Descriptive Statistics for Communication Subscale

Variables	*N* (%)	*M*	*SD*	Min	Max	Skew	Kurtosis
Gender							
Male	445 (50)	3.369	.857	.00	5.00	-.663	1.138
Female	443 (50)	3.482	.859	.67	5.00	-.551	-.004
Race/Ethnicity							
AA	98 (11)	3.498	.998	.00	5.00	-.844	1.005
Hispanic	222 (25)	3.319	.865	.50	5.00	-.631	.137
White	412 (46)	3.479	.848	.00	5.00	-.656	.849
Other	156 (18)	3.390	.778	1.42	5.00	-.155	-.213
SES							
Yes	455 (51)	3.316	.916	.00	5.00	-.554	.455
No	433 (49)	3.541	.782	.50	5.00	-.547	.414
Parent's Ed							
HS	408 (46)	3.248	.880	.50	5.00	-.379	-.125
SC	297 (33)	3.411	.872	.00	5.00	-.916	2.083
4yrs	183 (21)	3.843	.715	1.50	5.00	-.503	-.030
Native English							
Yes	716 (81)	3.429	.864	.00	5.00	-.589	.582
No	172 (19)	3.411	.843	.50	5.00	-.666	.530
Total	888	3.425	.860	.00	5.00	-.602	.565

No subgroup comprised less than 10% of the sample and no subgroup violated the assumption of normality. A casual examination of the mean values indicates potential

differences, but none larger than one Likert point. Some of the Kurtosis values were higher than expected, over 1.00, indicating that the data have some subgroups with many values in the tails of the distribution.

The descriptive statistics for the Interpretation subscale are presented in Table 4.

Table 4

Descriptive Statistics for Interpretation Subscale

Variables	*N (%)*	*M*	*SD*	Min	Max	Skew	Kurtosis
Gender							
Male	445 (50)	3.255	.906	.00	5.00	-.529	.683
Female	443 (50)	3.326	.859	.82	5.00	-.405	-.276
Race/Ethnicity							
AA	98 (11)	3.316	.957	.00	5.00	-.720	.915
Hispanic	222 (25)	3.214	.872	.00	5.00	-.722	.610
White	412 (46)	3.294	.900	.00	5.00	-.445	-.007
Other	156 (18)	3.377	.801	1.64	5.00	.126	-.497
SES							
Yes	455 (51)	3.218	.920	.00	5.00	-.515	.440
No	433 (49)	3.367	.837	1.00	5.00	-.376	-.145
Parent's Ed							
HS	408 (46)	3.186	.866	.00	5.00	-.406	.118
SC	297 (33)	3.233	.898	.00	5.00	-.620	.660
4yrs	183 (21)	3.620	.820	1.45	5.00	-.408	-.237
Native English							
Yes	716 (81)	3.277	.896	.00	5.00	-.464	.277
No	172 (19)	3.349	.828	1.00	5.00	-.514	.182
Total	888	3.291	.882	.00	5.00	-.477	.271

Once again, the subgroups sizes are the same and well within an acceptable range. The skewness values are well below the 1.00 mark, indicating that these data are approximately normal. There are no kurtosis values that are above the 1.00 mark, meaning that these data have a desirable distribution and are acceptable for analysis.

A review of the descriptive statistics for the problem formulation subscale reveals very similar results to that of the interpretation. The data are approximately normally distributed according to the skewness values and the data are not abnormally

concentrated in the tails or center, based on the review of the kurtosis values. The full descriptive results are found in Table 5.

Table 5

Descriptive Statistics for Problem Formulation Subscale

Variables	*N (%)*	*M*	*SD*	Min	Max	Skew	Kurtosis
Gender							
Male	445 (50)	3.438	.772	1.00	5.00	-.238	-.177
Female	443 (50)	3.488	.758	1.07	5.00	-.470	.016
Race/Ethnicity							
AA	98 (11)	3.428	.747	1.73	5.00	.005	-.415
Hispanic	222 (25)	3.339	.815	1.07	5.00	-.743	.190
White	412 (46)	3.535	.757	1.00	5.00	-.284	-.224
Other	156 (18)	3.472	.768	1.60	5.00	-.168	-.408
SES							
Yes	455 (51)	3.371	.774	1.00	5.00	-.271	-.185
No	433 (49)	3.560	.744	1.20	5.00	-.432	.062
Parent's Ed							
HS	408 (46)	3.335	.767	1.07	5.00	-.373	-.190
SC	297 (33)	3.443	.730	1.00	5.00	-.395	.035
4yrs	183 (21)	3.781	.727	1.47	5.00	-.301	-.334
Native English							
Yes	716 (81)	3.479	.766	1.00	5.00	-.246	-.306
No	172 (19)	3.398	.760	1.07	5.00	-.814	.629
Total	888	3.463	.765	1.00	5.00	-.351	-.103

The descriptive statistics for the Research subscale are presented in Table 6. Subgroup sizes did not differ for this subscale. The data were all approximately normally distributed with no skewness values higher than .7, well below the 1.00 mark which indicates a non-normally distributed data set. The data for one subgroup, African Americans, in this subscale had a kurtosis value of 1.259, which indicates that this distribution has many scores in the tails of the distribution, though this number is not too extreme.

Table 6

Descriptive Statistics for Research Subscale

Variables	*N (%)*	*M*	*SD*	Min	Max	Skew	Kurtosis
Gender							
Male	445 (50)	3.362	.810	.00	5.00	-.398	.503
Female	443 (50)	3.479	.809	1.00	5.00	-.510	.084
Race/Ethnicity							
AA	98 (11)	3.397	.927	.00	5.00	-.744	1.259
Hispanic	222 (25)	3.307	.800	.82	5.00	-.654	.388
White	412 (46)	3.476	.797	1.00	5.00	-.323	-.128
Other	156 (18)	3.450	.776	1.55	5.00	-.183	-.420
SES							
Yes	455 (51)	3.348	.846	.00	5.00	-.484	.451
No	433 (49)	3.496	.767	1.45	5.00	-.348	-.183
Parent's Ed							
HS	408 (46)	3.312	.813	.82	5.00	-.493	.211
SC	297 (33)	3.400	.782	.00	5.00	-.497	.606
4yrs	183 (21)	3.694	.795	1.45	5.00	-.384	-.160
Native English							
Yes	716 (81)	3.420	.816	.00	5.00	-.401	.236
No	172 (19)	3.621	.808	1.29	5.00	-.599	.176
Total	888	3.420	.811	.00	5.00	-.449	.268

And finally, Table 7 displays the descriptive statistics for the final subscale, precision and accuracy. The information is similar to the other descriptive for the four other subscales. The data are approximately normal with the highest skewness value of -.901. In addition, the African American subgroup, once again, had a kurtosis value of 1.350, indicating many values in the tails of the distribution. No other subgroups scores had kurtosis values above the 1.00 mark.

In general, the results of the descriptive statistics indicate that the data are cleaned, correctly coded, and meet some of the basic assumptions needed for further analyses.

Table 7

Descriptive Statistics for Precision & Accuracy Subscale

Variables	*N (%)*	*M*	*SD*	Min	Max	Skew	Kurtosis
Gender							
Male	445 (50)	3.474	.849	.00	5.00	-.414	.146
Female	443 (50)	3.694	.839	1.21	5.00	-.629	-.017
Race/Ethnicity							
AA	98 (11)	3.670	.936	.00	5.00	-.901	1.350
Hispanic	222 (25)	3.511	.833	1.14	5.00	-.646	.230
White	412 (46)	3.635	.862	1.00	5.00	-.497	-.284
Other	156 (18)	3.498	.779	1.71	5.00	-.095	-.564
SES							
Yes	455 (51)	3.475	.888	.00	5.00	-.458	-.002
No	433 (49)	3.698	.794	1.14	5.00	-.570	-.132
Parent's Ed							
HS	408 (46)	3.459	.850	1.21	5.00	-.386	-.332
SC	297 (33)	3.540	.840	.00	5.00	-.662	.631
4yrs	183 (21)	3.934	.776	1.64	5.00	-.604	-.085
Native English							
Yes	716 (81)	3.575	.861	.00	5.00	-.490	-.025
No	172 (19)	3.621	.808	1.29	5.00	-.599	.176
Total	888	3.584	.851	.00	5.00	-.510	.007

Correlation of Subscales

Because each of the five subscale variables was normally distributed and the assumption of linearity was not markedly violated, Pearson correlations were computed to examine the intercorrelations of the variables. Table 8 shows that the five variables were significantly and strongly correlated.

Table 8

Summary of Correlation Coefficients for Subscales

Subscale	2.	3.	4.	5.
1. Communication	.770**	.747**	.769**	.768**
2. Interpretation	--	.755**	.782**	.685**
3. Problem Formulation	--	--	.766**	.696**
4. Research	--	--	--	.739**
5. Precision & Accuracy	--	--	--	--

** $p < .01$

The strongest positive correlation, which would be considered a very large effect size according to Cohen (1988), was between the interpretation and research subscales, r (886) = .782, $p < .01$. This correlation means that students who had relatively high scores on the interpretation subscale were very likely to have high research subscale scores. The weakest correlation was between the interpretation and precision and accuracy subscales, although even these would be considered large effect sizes according to Cohen.

Analysis of Basic Differences

In order to answer the research questions concerning the basic differences between subgroups, independent-samples t-tests and a series of one-way ANOVAs were completed for each subgroup of interest. For all of the basic difference statistical tests, the alpha level was set at .05. What follows is a reporting of the findings for each of these subgroups; the findings of the independent samples *t*-tests are presented first.

Gender. In order to answer the question of whether differences exist in subscale scores between genders, an independent samples *t*-test was completed for each of the five subscales. Table 9 shows that males' scores were significantly different from females on three of the five subscales. Inspection of the two group means indicates that the average subscale scores for female students is significantly higher than males on communication (p = .049), research (p = .033), and precision/accuracy (p = .000). The effect sizes d ranged from .13 to .26, which are considered small effects. The evidence from the effect sizes was confirmed by an examination of means, which show little practical differences between the two groups, especially when considered on the original Likert scale.

Table 9

Comparison of Male and Female Students on Five Subscales

Variable	*M*	*SD*	*t*	*df*	*p*	*d*
Communication			-1.974	886	.049	.13
Males	3.369	.857				
Females	3.482	.859				
Interpretation			-1.200	886	.231	.08
Males	3.255	.906				
Females	3.327	.859				
Problem			-.980	886	.328	.07
Formulation	3.438	.772				
Males	3.488	.758				
Females						
Research			-2.139	886	.033	.14
Males	3.362	.810				
Females	3.479	.809				
Precision &			-3.891	886	.000	.26
Accuracy	3.474	.849				
Males	3.694	.839				
Females						

Economic status. The question of economic status is represented by the demographic item that asks students whether or not they qualify for free or reduced price meals (breakfast or lunch) at school (FARMS). Therefore, there are two groups for this question: those students who answered “yes” and those that answered “no”. Table 10 shows that students who do qualify for FARMS were significantly different from those who do not qualify on all subscales. Inspection of the two group means for each subscale indicates relatively small differences: .23 points for Communication, .15 points for Interpretation, .19 points for Problem Formulation, .15 for Research, and .22 points for Precision/Accuracy. In most cases, this would be equivalent to one point on the original Likert scale. The largest effect size was for the precision/accuracy subscale, although the effect would be considered small.

Table 10

Comparison of Students Who Do and Do Not Qualify for FARMS on Five Subscales (n = 455 "yes" and n = 433 "no")

Variable	*M*	*SD*	*t*	*df*	*p*	*d*
Communication			-3.950[A]	875.806[A]	.000	.26
Yes	3.316	.916				
No	3.541	.782				
Interpretation			-2.522	886	.012	.17
Yes	3.218	.920				
No	3.367	.837				
Problem			-3.711	886	.000	.25
Formulation	3.371	.774				
Yes	3.560	.744				
No						
Research			-2.733	886	.006	.18
Yes	3.348	.846				
No	3.500	.767				
Precision/Accuracy			-3.940[A]	882.613[A]	.000	.27
Yes	3.475	.888				
No	3.700	.794				

[A]The *t* and the *df* were adjusted because variances were not equal.

English language learner status. In order to answer the question of differences for students who consider themselves English language learners and those who do not, an independent samples *t*-test was completed. The item that identified the groups of this variable asked students "is English your first language". The equality of variance assumption was met for all subscales. A statistically significant difference was not found among the English language learners and their native speaking peers on any of the subscales. Table 11 reports the full findings of the independent samples *t*-test.

Table 11

Comparison of Native English Speaking and Non-native English Speaking Students on Five Subscales (n = 716 "yes" and n = 172 "no")

Variable	*M*	*SD*	*t*	*df*	*p*
Communication			.245	886	.716
Yes	3.4288	.864			
No	3.4109	.843			
Interpretation			-.962	886	.179
Yes	3.2769	.896			
No	3.3491	.828			
Problem Formulation			1.237	886	.394
Yes	3.4785	.766			
No	3.3982	.760			
Research			-.046	886	.411
Yes	3.4197	.816			
No	3.4229	.794			
Precision/Accuracy			-.641	886	.241
Yes	3.5750	.861			
No	3.6213	.808			

Parents' level of education. The one-way ANOVA was conducted to determine if the mean subscale score for each of five aspects differed on the level of education of the students' parents. The assumption of normality was tested and met for all subscales except for communication. According to Levene's test, the homogeneity of variance assumption was satisfied for the interpretation (F [2, 885] = .590, p = .555), problem formulation (F [2, 885] = .218, p = .804), research (F [2 885] = .020, p = .980), and precision/accuracy (F [2 885] = .844, p = .430) subscales only. The assumption of independence was also met.

The one-way ANOVA is statistically significant for all four subscales tested. Table 12 reports the results. The interpretation subscale difference was significant, F (2, 885) = 16.782 and p = .000. The problem formulation subscale difference was significant, F (2, 885) = 22.747 and p = .000. The research subscale difference was significant, F (2, 885) = 14.569 and p = .000.

Table 12

One-Way Analysis of Variance Summary Table Comparing Parents' Level of Educational Attainment on Five Subscales

Source	*df*	SS	*MS*	*F*	*p*
Interpretation				16.782	.000
Between	2	25.279	12.640		
groups	885	666.551	.753		
Within groups	887	691.830			
Total					
Problem Formulation				12.683	.000
Between	2	25.365	12.683		
groups	885	493.423	.558		
Within groups	887	518.788			
Total					
Research				9.299	.000
Between	2	18.598	9.299		
groups	885	564.894	.638		
Within groups	887	583.492			
Total					
Precision &				14.664	.000
Accuracy	2	29.328	14.644		
Between	885	612.527	.692		
groups	887	641.855			
Within groups					
Total					

And the precision/accuracy subscale difference was significant, F (2, 885) = 21.187 and p = .000. Because the communication subscale did not meet the assumption of homogeneity of variance, an alternative Welch procedure was used. The communication subscale difference was also found to be significant using this procedure, F_W (2, 500.915) = 38.575, p = .000.

A post hoc Hochberg's GT2 was conducted for each of the four significant differences in order to determine specifically where the differences can be found. Table 13 reports the full results of the post hoc testing. In all cases, the significant differences

were found between the "four years or more of college" group of students and the "some college" and "high school graduate or less" groups. Effect size, *d,* was calculated for each one of the post hoc comparisons. The largest effect size was found for the problem formulation subscale between students whose reported four years or more and those who reported high school or less, $d = .59$. The smallest effect size was between those who reported some college and four years or more on the research subscale.

Table 13

Summary of Results of Hochberg's GT2 Post Hoc Tests for Parents' Education Groups on Five Subscales

Variable	Group 1	Group 2	Mean Difference	SE	*p*	*d*
Interpretation	4 yr college or more	HS grad or less	.434	.077	.000	.51
		Some college	.387	.082	.000	.45
Problem Formulation	4 yr college or more	HS grad or less	.446	.066	.000	.51
		Some college	.339	.070	.000	.47
Research	4 yr college or more	HS grad or less	.382	.071	.000	.47
		Some college	.294	.075	.000	.37
Precision & Accuracy	4 yr college or more	HS grad or less	.475	.074	.000	.57
		Some college	.394	.078	.000	.48

Because at least two subscales did not meet the assumption of homogeneity of variance, a post hoc Games-Howell procedure was conducted in order to find the specific differences between groups. Similar results to the other post hoc tests were found. There was a larger than typical difference ($d = .72$) on the communication subscale between students who reported, "high school graduate or less" to the parent's education questions and the "four or more years of college" answer.

Race/Ethnicity. The question of differences between ethnic groups was answered using one-way ANOVA. Table 14 reports the results. Levene's test for equality of

variance was not significant for any of the subscales, indicating that the assumption of homogeneity of variance was not violated. A statistically significant difference was found for the problem formulation subscale only, $p = .021$. A Hochberg's GT2 post hoc test revealed that the significant difference for the problem formulation subscale is found between White and Hispanic students, with an effect size of $d = .25$, a small effect.

Table 14

One-Way Analysis of Variance Summary Table Comparing Ethnic Groups on Five Subscales

Source	*df*	SS	*MS*	*F*	*p*
Communication				1.984	.115
Between	3	4.385	1.462		
groups	884	651.351	.737		
Within groups	887	655.736			
Total					
Interpretation				1.074	.359
Between	3	2.512	.837		
groups	884	689.318	.780		
Within groups	887	691.830			
Total					
Problem Formulation				3.245	.021
Between	3	5.651	1.884		
groups	884	513.136	.580		
Within groups	887	518.788			
Total					
Research				2.185	.088
Between	3	4.296	1.432		
groups	884	579.196	.655		
Within groups	887	583.492			
Total					
Precision & Accuracy				1.917	.125
Between	3	4.148	1.383		
groups	884	637.707	.721		
Within groups	887	641.855			
Total					

Summary

The statistical procedures were run in three phases in order to address the specific research questions. All of the subscales were found to have a high correlation with one

another. The independent samples *t*-test and a series of one-way ANOVAs were run to determine if differences exist between groups of students on their mean subscale scores for each of the five subscales. Results for these difference tests were varied and indicated many small effects; however, a few large effects were found, indicating that there are some differences between groups of students. A more comprehensive discussion of the results of the study and the implications of those results follow in the next chapter.

CHAPTER V

DISCUSSION

The final chapter of this study is divided into three equally important sections. The first section provides a summary of the findings of the study, including a restatement of the problem and a summary of the methodology. The second section of the chapter explores practical implications of the study, situating it within the context of the literature review, theoretical model, and considerations for school practice. The final section points to the limitations of this study, potential areas for future research, and conclusions.

Summary of Findings

The purpose of this study, as stated in Chapter I, was to look for differences between the college readiness self-assessment subscale scores of subgroups of ninth and tenth grade students. Data from 888 ninth and tenth grade students from the Northwest of the United States were used to answer the research questions. The conceptual framework on which the study relies is Conley's Four Keys to College and Career Readiness model.

The methodology of this study, as outlined in Chapter III, was developed using a basic descriptive and inferential statistics design. First, a correlation analysis was conducted to determine if the subscales were statistically related. Then, a series of independent samples *t*-tests were conducted to examine basic differences between the gender, socioeconomic, and linguistic subgroups. Finally, analysis of variance tests, along with associated post hocs, were conducted to find differences between the ethnic and parental educational attainment subgroups.

Summary of results. The results of the study were outlined in Chapter IV. All five of the subscales are highly correlated. Of all of the subgroups examined in the study,

the most significant differences were found among students who reported their parents had completed four years or more of college and their peers who reported their parents had completed some college or high school/less than high school. These differences were found for all five subscales and all had medium to large effect sizes. Differences were also found among socioeconomic groups on the problem formulation and precision/accuracy subscales, though effect sizes were small. A small difference was found between White and Hispanic students on the problem formulation subscale and between male and female students on the precision/accuracy subscale only. In this study, no difference was found between English Language Learners and their native-English speaking peers on any of the subscales.

Implications of the Study

As noted in Chapter IV and the brief summary of findings in the previous section of this chapter, results of this study were mixed, with differences found for some groups of students and on only some of the subscales. The following section explores the implications of the study for each of the findings.

Findings related to subscale association. The results of the correlation analysis indicate that all five of the subscales were highly correlated. It is important to recognize that this correlation is not so high as to indicate that the subscales are each measuring the exact same construct. For example, extremely high correlation coefficients, about .9 range, might indicate that the subscales are measuring the same thing, in which case it would not be appropriate to give respondents all five subscales. This was not the case in this study; instead, correlation coefficients indicate that the subscales, while conceptually related, are also statistically related.

Findings related to subgroup identification. According to the results of this study, the subgroups to which a student belongs can have an effect on a student's self-assessment of his/her college readiness. Though the differences are varied, both in their magnitude and the number of subscales in which they can be found, the findings of this study provide some confirmation of the inequities that exist in American college preparedness.

Parental educational attainment. The strongest evidence of differences between groups of students on the five subscales was found for the question of parental educational attainment. The findings indicate that, on average, students whose parents had completed four years or more of college education had higher subscale scores than their peers. These findings reinforce research, which has shown that students who might be considered first-generation college goers face a particular challenge when it comes to preparing for success in postsecondary education (Venezia & Kirst, 2005). Given that these students' self-assessments are already indicating a difference at the ninth and tenth grade, it seems appropriate that individualized interventions be developed at earlier grades to help students gain more skill with the cognitive strategies necessary for success.

Socioeconomic status. Some research has found that socioeconomic status may have an effect on the college readiness of students; for example, Dahlin & Tarasawa (2013) found that students in low-poverty schools were better prepared for college than their peers in high-poverty schools, and that this difference between the two groups remains essentially constant throughout the years that these students were tracked. Additionally, only 52% of low-income students were estimated to have immediately enrolled in college in 2004, while 80% of high-income students enrolled in the same year

(Roderick et al., 2009). The second largest number of significant differences in this study was among students who qualify for free and reduced meals at school and their peers who do not. These students differed on the problem formulation subscale and the precision/accuracy subscale. On average, students who did not qualify for FARMS had higher self-assessment subscale scores on problem formulation and precision/accuracy. The fact that these students have a higher opinion of their own college readiness appears to be in line with other predictors of readiness, which show that these students are more prepared for college than their peers who fall in the lower socio-economic groups.

Race/Ethnicity. Students of color, along with students of low-income backgrounds, continue to bear the brunt of inequitable educational practices and some research has shown that readiness for college correlates with racial and socioeconomic inequality (Castro, 2013). In 2004, only 58% of Hispanic students are estimated to have enrolled immediately in college, compared to almost 70% of their White peers (Roderick et al., 2009). There is a clear imperative to deal with issues regarding race/ethnicity as they pertain to college readiness. This study only found one statistically significant difference on one of the five subscales among ethnic groups. A small difference exists between White and Hispanic students on their self-assessments for the problem formulation subscale. Based on their responses, these students appear to believe that the items measuring each of the Key Cognitive strategies are "somewhat like" themselves. These findings stand in contrast to a larger body of evidence in the scholarship that race and ethnicity can have large effects on a student's college preparedness (Roderick et al., 2004; Cates & Shaefle, 2012; Castro, 2013). Additional questions arise from these findings (see later section on future research).

Gender. There are clear gender differences in many traditional predictors of college readiness, including standardized assessment scores and academic course history. In fact, Combs et al. (2010) found that boys scored marginally higher than girls on standardized assessment indicators of readiness such as the SAT and ACT.

In this study, females had higher subscale scores than males on three of the five subscales. Though statistical differences were found in this study, from a practical perspective, they were very small. The subscale with the largest difference was precision/accuracy, where females had higher subscale scores than males by .3 of a point. No research evidence suggests that females are naturally more precise than their male peers, but perhaps they are more likely to assess their own skills more accurately than males. This study adds some potential support to an argument advanced by Hattie (2009) that differences between genders in most school settings are generally overstated and that where differences exist, they are typically very small.

ELL status. As stated in the review of the research literature in Chapter II of this study, the research around college readiness of students for whom English is not a first language is limited. The results of this study show that there were no significant differences between native English speakers and their non-native speaking peers. There is, however, some evidence in the limited body of research that fewer English language-learning students are entering college when compared to their native-speaking peers (Kanno & Cromley, 2013). There is also some evidence that ELLs face other challenges in K-12 education contexts, such as poor achievement outcomes on measures of academic progress. Certainly, this is an area that will need further research into the college readiness of this important subgroup. Perhaps by focusing on supporting those

students to access rigorous, comprehensive, and high-leverage courses, especially those marked as being predictive of success in college, schools can be sure that these students are better prepared.

Findings related to school-based practice. Public education should be a democratizing force in American society. Educational research that highlights persistent inequities in the system force school-based practitioners to consider new ways to prepare students. After all, a student's preparedness for college should not be determined by things outside of his/her control, like race and ethnicity, socioeconomic status, or parent's level of educational attainment. The findings of this study, which confirm the presence of some differences, add support to that larger conversation and can have the following broader implications.

The first implication of the findings of this study is that secondary schools, in the sample area of the in the United States, must continue to develop courses that incorporate cognitive strategies in their programs of study. Critically, these courses and their instructors should be explicit in labeling the "ways of thinking" that are necessary for college success. As students become more familiar with these terms and their use, they may be better able to assess their own skills.

Recent reform in K-12 education that has included the adoption of the Common Core State Standards has allowed for schools to pivot to curricula that are more closely aligned with college-level expectations and which include more cognitively challenging activities (Conley et al., 2011). The findings of this study certainly imply that students can benefit from additional practice with academic tasks, which require key cognitive strategies. Examples of these types of tasks include opportunities for students to

examine/evaluate models, create stylistic imitations, and present their products in effective ways (Conley, 2012).

A second implication of the findings of this study is that students must be given equitable access to the courses, which provide these high-leverage academic experiences. Schools should avoid inequitable practices, such as tracking of students that prevent them from having access to courses that provide opportunity to develop cognitive strategies. Though the data from this study do not point to whether or not students have taken specific courses, the broader research literature notes that Hispanic students and low-income students are less likely to have access to advanced courses, particularly in math and science, which may include cognitively challenging assignments (Castro, 2013).

Finally, given the differences among groups of students whose parents have or have not attended college, schools can play a role in providing information and messaging about college. For example, Venezia & Kirst (2005) found that teachers could play a vitally important role in educating students about academic expectations for college, procedures for applying to college, and ways to improve preparedness for college. If parents do not have the personal experience or external resources providing them information, it is imperative that schools fill that void to provide that information to all groups of students.

Limitations of the Study

Despite efforts to ameliorate some of the potential limitations of this study, there are constraints that threaten both the internal and external validity of the interpretations and findings. What follows is a brief examination of some of those issues.

First, it must be acknowledged that this study was a secondary analysis of data already collected for a slightly different purpose. Therefore, it was not possible to control some of the conditions that can have an effect on results. One example of this is the lack of control over administration procedures at each of the participating school sites. Though EPIC did provide clear directions and coaching to participating schools on the ways to identify a random sample and the important administration procedures, it is not possible to verify if each site followed these directions with fidelity. Implementation fidelity is a crucial consideration when addressing internal validity. The issue of whether or not the sample is truly random also limits the generalizability of results of this study. It is not possible to generalize the results to any population beyond that which was sampled for the study.

Another limitation of the study could be the use of self-assessment. While there are many advantages to self-assessment (see Chapter II), at least two disadvantages should be considered in the measurement of a construct. Chief among the disadvantages is the fact that self-assessment is subjective and allows for students to over or underestimate their own abilities. Additionally, students, for a variety of reasons, may inaccurately report some of the demographic variables. When designing this study, the possibility of inaccurate reporting was considered, and demographic variables that were less likely to be incorrectly reported (e.g. gender, ethnicity, etc.) were selected.

Additionally, the deletion of some of the data could pose a problem to the validity of inferences made. Approximately 29% of the total sample of ninth and tenth graders eligible for inclusion in this study was deleted from the sample because of a significant amount of missing data. For example, some of the students did not answer any of the

questions for the five subscales of interest and others did not answer enough of the items for the subscales to be included. The most common reason for data to be missing was due to item nonresponse. In order to delete respondents with missing data, the procedure known as *list wise deletion* was chosen. Though this procedure is common practice in the research literature, it does introduce potential threats to validity. The major limitation of this approach is the possibility that the missing data indicate different inferences than the complete data. Therefore, an analysis of only the complete data runs the risk of being biased. A potential way to mediate this limitation would be to produce an analysis that indicates that the data that are missing are missing completely at random. This further step in analysis would be an excellent opportunity for future research.

Finally, some methodologists may question the use of Likert variables in this study and their treatment as approximately continuous variables. There is a lively discussion among research methodologists about the use of Likert variables and the statistical tests, which may be performed using them. As Norman (2010) advances, Likert data have been debated in terms of whether or not they may be considered ordinal or interval scale, a decision which greatly affects the type of statistical analysis that can be conducted using them. According to Norman, these debates often overlook the fact that modern parametric statistical methods (like ANOVA) are robust and are less affected by violations of assumptions than their older, more simplistic predecessors (p. 3). In order to ascribe to Norman's, and by extension the researchers he cites, point of view on this topic, this study uses a manipulation of the variables. By averaging the scores of items to create each subscale, the Likert scores have been transformed to approximately continuous variables thereby lessening the potential questions caused by their use.

Future Areas of Research

Though the findings of this study do point to some differences that exist between subgroups of students, additional research is necessary to untangle the complicated issues surrounding the research questions. More specifically, future research could explore (a) the self-assessment scores of these students on subscales for the other three dimensions of Conley's Four Keys model, (b) the perceptions of teachers, administrators, and counselors for the subscales of interest, (c) the students' level of understanding of the terms used in the self-assessment tool, (d) comparison of self-assessment scores for key cognitive strategies with other measures of the same.

Adding additional data to the study that could compare/contrast student self-assessment scores with other traditional predictors of college readiness would possibly add interesting information to the study. In fact, the opportunity to triangulate the results of the self-assessment with other related data would provide a guaranteed complexity that might point to slightly different results. For example, it might add value to compare self-assessment scores with grade point average, SAT/ACT scores, or academic history. This information would need to be collected from the schools or districts rather than be self-reported by students. Additionally, the weaknesses of these traditional methods of predicting readiness would need to be acknowledged and/or neutralized if possible.

Another interesting area of research would be to examine the students' self-assessment scores alongside those of their teachers. For example, one could consider if the students' self-assessment matches the assessment of the appropriate teacher for that skill. Similarly, one could consider if students provide confirmatory evidence for their teachers. For example, if a teacher says that her assignments frequently require students

to problem solve, do students agree with her statement? Of course, in order to be able to make these comparisons, it would be important to match students and teachers.

The areas for future research listed in the section above are not intended to be an exhaustive list. There are many avenues of research that should be and are being pursued by researchers and practitioners in the field of college and career readiness.

Conclusions

From this study, two general conclusions can be drawn. From the subgroup identity variables examined in this study, the "level of educational attainment of parents" variable consistently indicated a difference on the five self-assessment subscales; thus, for this sample, there was a significant effect of parent's level of educational attainment on subscale scores. These differences were also the largest differences according to the effect size measures. The second general conclusion is that the socioeconomic status variable (whether or not a student qualifies for FARMS) indicated small differences on the subscales. The conclusions for the other research questions concerning differences for other subgroups are inconclusive. While small differences were found for gender and race/ethnicity, no significant difference was found for English language learner status, which is an important finding in its own right and merits further research.

REFERENCES CITED

ACT, Inc. (2004). *Crisis at the core: preparing all students for college and work*. Retrieved from www.act.org

ACT, Inc. (2007). *The role of nonacademic factors in college readiness and success*. Retrieved from www.act.org

ACT, Inc. (2012). *Measuring college and career readiness: The class of 2009*. Retrieved from www.act.org

ACT, Inc. (2013). *The condition of college & career readiness 2013*. Retrieved from www.act.org

Adelman, C. (2006). The Toolbox Revisited: Paths to Degree Completion From High School Through College. Washington, D.C.: U.S. Department of Education.

Allen, J. (2013). *Updating the ACT college readiness benchmarks* (Report No. 2013-6). Iowa City: ACT, Inc. Retrieved from www.act.org

Allen, J. & Sconing, J. (2005). Using ACT Assessment Scores to Set Benchmarks for College Readiness. *ACT Research Report Series (*Report No. 2005-3). Retrieved from www.act.org

Andrade, H. & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education, 2*, 159-181. doi: 10.1080/02602930600801928

Athanasou, J. A. (2012). Adult language, literacy, numeracy, and problem-solving skills in the workplace. Australian Journal of Adult Learning, 52, 173-182.

Attewell, P., Lavin, D., Thurston, D. & Levey, T. (2006). New evidence on college remediation. The Journal of Higher Education, 77, 886-924.

Barnes, W., Slate, J., & Rojas-LeBouef, A. (2010). College-readiness and academic preparedness: The same concepts? Current Issues in Education, 13. Retrieved from http://cie.asu.edu/

Brown, S. D., Tramayne, S., Hoxha, D., Telander, K., Fan, X., & Lent, R. W. (2008). Social cognitive predictors of college students' academic performance and persistence: A meta-analytic path analysis. Journal of Vocational Behavior, 72, 298-308. doi:10.1016/j.jvb.2007.09.003

Castro, E. L. (2013). Racialized readiness for college and career: Toward an equity-grounded social science of intervention programming. *Community College Review, 41*, 292-310. doi: 10.1177/0091552113504291

Cates, J. T. & Schaefle, S. E. (2011). The relationship between a college preparation program and at-risk students' college readiness. Journal of Latinos and Education, 10, 320-334.

Chang, C., Liang, C., & Chen, Y. (2012). Is learner self-assessment reliable and valid in a web-based portfolio environment for high school students? *Computers & Education, 60*, 325-334.

Cimetta, A. D., D'Agostino, J. V., & Levin, J. R. (2010). Can high school achievement tests serve to select college students? Educational Measurement: Issues and Practice, 29, 3-12.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* New Jersey: Lawrence Erlbaum Associates.

College Board (2012). 2012 College-Bound Seniors Total Group Profile Report. Retrieved from www.media.collegeboard.com

College Board (2011). The college completion agenda: Progress report 2011 Retrieved from http://www.completiongagenda.collegeboard.org

Combs, J. P., Slate, J. R., Moore, G. W., Bustamente, R. M., Onwuegbuzie, A. J., & Edmonson, S. L. (2010). Gender differences in college preparedness: A statewide study. Urban Review, 42, 441-457.

Complete College America (2012). Remediation: Higher education's bridge to nowhere. Retrieved from http://www.completecollegeamerica.org

Conley, D. T. (2007). Toward a more comprehensive conception of college readiness. Eugene, OR: Educational Policy Improvement Center.

Conley, D. T. (2008). Rethinking college readiness. *New directions for higher education, 144,* 3-13. doi: 10.1002/he

Conley, D. T. (2011). Building on the common core. *What Students Need to Learn*, *68*, 16-20.

Conley, D. T. (2012a, March). *College and Career Readiness for ALL Students: More than a Cut Score.* In *Transformation in Public Education*. Symposium conducted at the meeting of the Oregon Education Association, Wilsonville, OR.

Conley, D. T. (2012b). *A Complete Definition of College and Career Readiness*. Educational Policy Improvement Center.

Conley, D. T. (2013). Rethinking the notion of 'noncognitive'. *Education Week,* 1-3.

Conley, D. T., McGaughy, C., Kirtner, J., van der Valk, A., & Martinez-Wenzl, M. T. (2010, April). College readiness practices at 38 high schools and the development of the college Careerready school diagnostic tool. Paper presented at the meeting of the American Educational Research Association, Denver.

Dahlin, M. & Tarasawa, B. (2013). *A level playing field? How college readiness standards change the accountability game.* Portland: Northwest Evaluation Association.

Dunning, D., Heath, C., & Suls, J. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 3,* 69-106.

Dzubak, C. M. (2010). The cognition gap: Sufficient skills for high schools but not sufficient for college. Synergy, 3, 1-20. Retrieved from http://myatp.org.

Educational Policy Improvement Center (EPIC) (2013). www.epiconline.org

Falchikov, N. & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 49.4*, pp. 395-430.

Field, A. (4th ed). (2013). *Discovering statistics using IBM SPSS statistics*. Los Angeles: Sage.

Ganzert, B. (2012). The effects of dual enrollment credit on gender and race. *Current Issues in Education, 15,* 1-9.

Greene, J. P. & Forster, G. (2003). Public High School Graduation and College Readiness Rates in the United States (Report No. 3). New York: Manhattan Institute for Policy Research.

Hahs-Vaughn, D. (2004). The impact of parents' education level on college students: An analysis using the beginning postsecondary students longitudinal study 1990-92/94. Journal of College Student Development, 45, 483-500. doi: 10.1353/csd.2004.0057

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* New York: Routledge.

Heritage, M. (2009). Using self-assessment to chart students' paths. *Middle School Journal, 40.5,* pp. 27-30.

Ibabe, I. & Jauregizar, J. (2010). Online self-assessment with feedback and metacognitive knowledge. *Higher Education, 59*, pp. 243-258. doi: 10.1007/s10734-009-9245-6

Kallison, J. M. and Stader, D. L. (2012). Effectiveness of summer bridge programs in enhancing college readiness. Community College Journal of Research and Practice, 36, 340-357.

Kanno, Y. & Cromley, J. G. (2013). English language learners' access to and attainment in postsecondary education. *TESOL quarterly, 47,* 89-121.

Kirby, N. & Downs, C. (2007). Self-assessment and the disadvantaged student: Potential for encouraging self-regulated learning. *Assessment & Evaluation in Higher Education, 4*, pp. 475-494. doi: 10.1080/02602930600896464

Kirst, M. (2004). The high school/college disconnect. *Educational leadership,* 51-55.

Knox, A. (2006). "Why American business demands twenty-first century learning: A company perspective." New Directions for Youth Development, 110, 31-37.

Kobrin, J. L. (2007). *Determining SAT benchmarks for college readiness* (Report No. RN-30). New York: The College Board.

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). Validity of the SAT for Predicting First-Year College Grade Point Average (Report No. 2008-5). New York: The College Board.

Komarraju, M., Ramsey, A., & Rinella, V. (2013). Cognitive and non-cognitive predictors of college readiness and performance: Role of academic discipline. Learning and Individual Differences, 1-7.

Lawrence, I. M., Rigol, G. W., Van Essen, T., & Jackson, C. A. (2003). *A historical perspective on the content of the SAT*. (Report No. 2003-3). New York: The College Board.

Lombardi, A., Seburn, M., & Conley, D. (2011). Development and initial validation of a measure of academic behaviors associated with college and career readiness. Journal of Career Assessment, 19,375-391. doi: 10.1177/1069072711409345

Lombardi, A. R., Conley, D.T., Seburn, M. A. , Downs, A.M. (2013). College and career readiness assessment: Validation of the key cognitive strategies framework. *Assessment for effective intervention, 38,* 163-71. doi: 10.1177/1534508412448668

Long, M. C., Iatarola, P., & Conger, D. (2009). Explaining Gaps in Readiness for College-Level Math: The Role of High School Courses. Education Finance And Policy, 4(1), 1-33.

Marcum, J. (2004). Information literacy. In A. Distefano, K. Rudestam, & R. Silverman (Eds.), Encyclopedia of distributed learning. (pp. 228-232). Thousand Oaks, CA: SAGE Publications. doi: 10.4135/9781412950596.n81

Maruyama, G. (2012). Assessing college readiness: Should we be satisfied with ACT or other threshold scores? Educational Researcher, 41, 252-261. doi: 10.3102/0013189X12455095

Moore, G. W., Slate, J. R., Edmonson, S. L., Combs, J. P., Bustamante, R., & Onwuegbuzie, A.J. (2010). High school students and their lack of preparedness for college: A statewide study. *Education and Urban Society*, 42, 817-838. doi: 10.1177/0013124510379619

Nankervis, B. (2011). Gender Inequities in University Admission due to the Differential Validity of the SAT. Journal Of College Admission, (213), 24-30.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in health science education,* 1-8. doi:10.1007/s10459-010-9222

O'Connor, L. G., Radcliff, C. J., & Gedeon, J. A. (2002). Applying systems design and item response theory to the problem of measuring information literacy skills. College and Research Libraries, 63, 528-43.

Partnership for 21st Century Skills. (2011). Framework for 21st century learning. Washington DC: P21.

Patterson, B. F. & Mattern, K. D. (2013). Validity of the SAT for Predicting First-Year Grades: 2010 SAT Validity Sample (Report No. 2013-2). New York: The College Board.

Pattison, E., Grodsky, E., & Muller, C. (2013). Is the sky falling? Grade inflation and the signaling power of grades. *Educational Researcher, XX,* 1-7. doi: 10.3102/0013189X13481382

Perez, C. (2002). Different tests, same flaws: Examining the SAT I, SAT II, and ACT. *The Journal of College Admission, Fall*, 20-25.

Porter, A., McMaken, J., Hwang, J., & Yang, R. (2011) Assessing the common core standards: Opportunities for improving measures of instruction. Educational Researcher, 40, 186-194.

Porter, A. C., & Polikoff, M. S. (2012). Measuring academic readiness for college. Educational Policy, 26, 396-411. doi: 10.1177/0895904811400410

Radunzel, J. & Noble, J. (2012). Predicting Long-Term College Success through Degree Completion Using ACT Composite Score, ACT Benchmarks, and High School Grade Point Average. Research Report No. 2012-5. Retrieved from www.act.org

Robbins, S. B., Allen, J., Casillas, A., Hamme Peterson, C., & Le, H. (2006). Unraveling the differential effects of motivational and skills, social, and self-management measures from traditional predictors of college outcomes. *Journal of Educational Psychology, 98*, 598-616. doi: 10.1037/0022-0663.98.3.598

Roderick, M., Nagaoka, J., & Coca, V. (2004). College readiness for all: the challenge for urban high schools. *The Future of children, 19,* 185-211.

Rothman, R. (2012). A common core of readiness. *Educational leadership, April,* 11-18.

Saiz, C. & Rivas, S. F. (2011). Evaluation of the ARDESOS program: An initiative to improve critical thinking skills. *Journal of Scholarship of Teaching and Learning, 2,* 34-51.

Shook Cheong, A. (2005). Problem solving. In C. Fisher, & R. Lerner (Eds.), Encyclopedia of applied developmental science. (pp. 879-880). Thousand Oaks, CA: SAGE Publications. doi: 10.4135/978141295065.n337

Smyth, J. & Terry, C. (2007). "Self-Report" in *Encyclopedia of Measurement and Statistics.* Ed. Neil Salkind. Thousand Oaks: Sage

Snyder, T. D., Dillow, S. A., and Hoffman, C. M. (2009). Digest of Education Statistics 2008 (NCES 2009-020). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Stemler, S. E. (2012). What should university admissions tests predict? *Educational Psychologist, 47*, 5-17. doi: 10.1080/00461520.2011.611444

Sternberg, R. J. (2006). The rainbow project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence, 34*, 321-350.

Sparks, D. & Malkus, N. (2013). Statistics in brief: First-year undergraduate remedial coursetaking—1999-2000, 2003-04, 2007-08. (NCES No. 2013-013). Retrieved from National Center for Education Statistics website: http://www.nces.org

Toothaker, L.E. (1993). *Multiple comparison procedures.* Sage University paper series on quantitative applications in the social sciences. Newbury Park: Sage.

Van Merrienboer, J. J. G. (2013). Perspectives on problem solving and instruction. *Computers & Education, 64*, pp. 153-160.

Venezia, A. & KIrst, M. W. (2005). Inequitable opportunities: how current education systems and policies undermine the chances for student persistence and success in college. Educational Policy, 19, 283. doi: 10.1177/0895904804274054

Wagner, T. (2010). *The global achievement gap: Why even our best schools don't teach the new survival skills our children need—and what we can do about it.* New York: Basic.

Wilkins, C., Rolfhus, E., Hartman, J., Brasiel, S., Brite, J., Howland, N., & Regional Educational Laboratory Southwest (2012). How Prepared Are Subgroups of Texas Students for College-Level Reading: Applying a Lexile[R]-Based Approach. REL Technical Brief. REL 2012-No. 018. Regional Educational Laboratory Southwest.

Wyatt, J., Kobrin, J., Wiley, A., Camara, W. J., & Proestler, N. (2011). *Development of a college readiness benchmark and its relationship to secondary and postsecondary school performance* (Report No. 2011-5). New York: The College Board.

Young, J. W. (2003). The past, present, and future of the SAT: Implications for college admissions. *Journal of the American Association of Collegiate Registrars and Admissions Officers, 78*, 21-24.

Yun, J. T. & Moreno, J.F. (2013). College access, k-12 concentrated disadvantage, and the next 25 years of education research. *Educational Researcher, 35,* 12-21. doi: 10.3102/0013189X035001012