SUPPLEMENTING HONORS COLLEGE APPLICATIONS WITH AN OPTIONAL ESSAY: EFFECTS ON COHORT DIFFERENCES, FIRST YEAR RETENTION, AND ETHNIC AND GENDER DISPARITIES

By

MCKENZIE LEIGH MOHLER

Bachelor of Arts in English Literary Studies

Washington State University

Pullman, Washington

2010

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December 2013

SUPPLEMENTING HONORS COLLEGE APPLICATIONS WITH AN OPTIONAL ESSAY: EFFECTS ON COHORT DIFFERENCES, FIRST YEAR RETENTION, AND ETHNIC AND GENDER DISPARITIES

Thesis Approved:

Dr. Tami Moore

Thesis Adviser

Dr. Jesse Mendez

Dr. Robert Spurrier

Name: MCKENZIE LEIGH MOHLER

Date of Degree: DECEMBER, 2013

Title of Study: SUPPLEMENTING HONORS COLLEGE APPLICATIONS WITH AN OPTIONAL ESSAY: EFFECTS ON COHORT DIFFERENCES, FIRST YEAR RETENTION, AND ETHNIC AND GENDER DISPARITIES

Major Field: EDUCATIONAL LEADERSHIP STUDIES, OPTION IN HIGHER EDUCATION ADMINISTRATION

Abstract:

The purpose of the present study was to examine potential differences in background characteristics and academic achievement of honors students who submitted different types of applications. Students who applied to an honors college using traditional applications (high school grade point average and standardized test scores) were compared to students who applied to the honors college using augmentative applications (adding an optional essay component to the traditional application). Nine hundred and eight six students in two consecutive incoming classes were included in this study, which blended causal-comparative research and correlational research. Sternberg's WICS model was the theoretical framework for the study. Independent variables included matriculation year, type of application submitted, and fall-to-fall retention. Dependent variables included fall-to-fall retention, first-year cumulative grade point average, race or ethnicity, gender, reported parental income, standardized test scores, and high school grade point average.

Findings indicated that students admitted in the second cohort when augmentative essays were available only differed from the first cohort in race or ethnicity; the percentage of non-white students significantly increased among the second cohort. Students who submitted augmentative essays when they already met the minimum requirements to join the honors program had lower ACT scores than students who also met the minimum requirements and did not submit an augmentative application. Students who submitted augmentative essays when they did not meet the minimum requirements to join the honors program were less likely to be retained one year later than students who met the minimum requirements, and they had lower first year grade point averages. High school grade point average was the strongest predictor of both fall-to-fall retention and first year grade point average. Recommendations included decreasing reliance on ACT scores in honors admissions and considering other alternative admissions tools to increase honors program student diversity.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION TO THE STUDY

Background of the Study

University administrators founded honors programs and colleges (hereafter collectively referred to as programs, though the distinction can be significant [National Collegiate Honors Council, 2010a; National Collegiate Honors Council, 2010b]), to provide a uniquely individualized experience for high-talent undergraduate students (Long, 2002). Benefits of participating in honors programs include increased graduation rates, increased grade point averages, and shorter time to degree (Cosgrove, 2004). Honors programs at less expensive and prestigious institutions can be seen as a way to enhance the level of education available to students from lower socioeconomic backgrounds.

Honors programs, like institutional admissions offices, utilize numerous criteria to sift through undergraduate applications and select for admission those students they find qualified and believe capable of persisting and graduating (Cosgrove, 2004; McKay, 2009; Moon 2012). Institutions use standardized test scores, usually the SAT (initially the Scholastic Aptitude Test, now not considered an acronym) or ACT (originally the American College Test, also no longer an acronym), to equalize applicants, because factors such as grade point averages, community involvement, and letters of recommendation can vary widely in significance for different schools, districts, and states (Epenshade, Hale, & Chung, 2005; Sternberg, 2010b; Zwick, 2007). Admissions offices use the SAT and ACT because of the tests' ability to predict students' first year grade point averages (FYGPA); when standardized test scores are combined with high

school grade point averages (HSGPA), the predictive strength of both increases (Geiser & Studley, 2002). Scholars and critics have raised two distinct objections to the use of standardized tests in admissions: First, they may not predict FYGPA and eventual college success as well as is purported; secondly, research has shown that ethnic minority groups, especially blacks and Latinos, and students from low socioeconomic backgrounds, do not score as well on the tests as their white peers and students from high socioeconomic backgrounds (Atkinson & Geiser, 2009; Rothstein, 2004). Admitting ethnically and socioeconomically diverse students to college is important because meaningful interaction with peers from different racial or ethnic groups results in both positive educational outcomes, such as active thinking and intellectual engagement, and prepares students to participate in a diverse democratic society with outcomes such as citizenship engagement, the ability to see things from others' perspectives, and racial/cultural engagement (Gurin, Dey, Hurtado, & Gurin, 2002). After undertaking a series of studies on standardized tests regarding their impact on non-white and white students from low socioeconomic backgrounds, Richard Atkinson, then president of The University of California system, made a proposal for his institutions to stop requiring the SAT or ACT for admission in 2001 (Atkinson & Geiser, 2009). In the wake of this movement, alternative or augmentative (Sternberg, 2010a) admissions criteria have been presented to cultivate more diverse and equitable college cohorts. These alterations to college admission policies come at a time when the use of affirmative action in higher education has also been questioned or legally prohibited (Gratz v. Bollinger, 2003; Hopwood v. Texas, 1996; Regents of the University of California v. Bakke, 1978), leaving institutions to seek new ways to recruit and retain individuals of different ethnicities, religions, socioeconomic statuses, sexual orientations, and intelligence types (McKay, 2009; Rothstein, 2004; Sternberg, 2010b). The democratic goal of open access to higher education necessitates an understanding of alternative methods to admission beyond standardized tests, which may reinforce socioeconomic barriers to higher education and contribute to the increasing income and opportunity disparities

between socioeconomic and ethnic groups (Atkinson & Geiser, 2009; Mohler, 2012b; Rothstein, 2004; Sternberg, 2010b).

While the efforts of institutions to recruit and admit diverse students are closely examined by the public and by scholars, little research has been done on increasing diversity in honors programs through alternative admissions policies. While the importance of diversity in honors programs in many respects reflects the importance of diversity in universities overall, there are additional aspects increasing the exigency of admitting ethnically and socioeconomically diverse students to honors programs. Participation in honors benefits all students in terms of learning outcomes such as course-related peer interaction, instructor skill and feedback, interaction with faculty, and cognitive development, but students of color and men experience even greater growth in learning outcomes from honors participation than white students and females (Seifert, Pascarella, Colangelo, & Assouline, 2007; Shushok, 2002). In light of these potential benefits, honors educators have declared:

> If honors programs and colleges are not serving those from the lower class, from racial and ethnic minority groups, from rural areas, from non-English-speaking backgrounds, and so on, it is our responsibility to make sure these students are identified and provided the appropriate education for their ability level. Honors education involves taking the abilities or potential abilities of an intellectually advanced group and nurturing those as much as possible, just as we would nurture a group of students at any level of intellectual ability, high or low. This is not 'elitism'; this is providing equal opportunity. (Rinn & Cobane, 2009)

In all honors programs, diversity is an issue that should receive attention, particularly in recruitment and admission. This quantitative study to examined the use of augmentative admission essays in one honors college and subsequent impacts on student diversity and retention. The research paper will include five main sections: the introduction, the literature review, the methodology, the findings, and the conclusion. The introduction will include the

purpose of the study, an overview of the methodology, assumptions and limitations, and the significance of the proposed study. The literature review will describe relevant research, including Sternberg's (2005) WICS (Wisdom, Intelligence, Creativity, Synthesized) Model and augmented theory of successful intelligence (Sternberg, The Rainbow Project Collaborators, and the University of Michigan Business School Project Collaborators, 2004), which provides the theoretical perspective for the study. The methodology will describe the population, selection of participants, procedures, and intended method of data analysis. The findings section will discuss the statistical tests run, presented for each of the two research questions, grouped by independent variable and then by dependent variable. The conclusion will discuss the results of the statistical analysis, including interpretation of the findings, relationship to prior research, and implications and recommendations for further research, the theory, and practice. This research adds to the paucity of quantitative studies regarding admissions-driven efforts to increase diversity in honors programs, and contributes understanding to Sternberg's WICS model (2005) and Kaleidoscope project (2010b) in new populations.

Purpose of the Study

The purpose of the present study was to examine potential differences in background characteristics and academic achievement of honors students who submitted different types of applications. Students who applied to the honors college using traditional applications (HSGPA and standardized test scores) were compared to students who applied to the honors college using augmentative applications (adding an optional essay component to the traditional application). Students apply to the honors college in an entirely separate process from their application to the university overall; the honors college cannot admit students until after the university has admitted them, and the honors application in no way impacts students' university applications. Research questions for the study included:

1. Do students who submitted augmentative applications differ in race or ethnicity, socioeconomic status, or academic achievement from students who submitted traditional applications?

2. Are students who submitted augmentative applications retained differently than students who submitted traditional applications?

It was hypothesized that students who submitted augmentative applications would differ from students who submitted traditional applications, and that augmentative applicants would be retained differently than students who submitted traditional applications. The participants were analyzed on multiple dependent variables, including fall-to-fall retention, first-year cumulative grade point average, race or ethnicity, gender, reported parental income, standardized test scores, and high school grade point average. Standardized test scores were converted to a single scale, ACT scores, for comparison. An honors college at a large four-year public land-grant research institution (Carnegie Foundation for the Advancement of Teaching, 2010) in the Midwest was the location of study.

Assumptions and Limitations

Based on Sternberg's (2010b) findings from prior research, an underlying assumption was that those students who wrote the optional essay might have been more motivated or enthusiastic about joining The Honors College. However, testing this assumption was outside the scope of this study.

The assumption that the ACT and SAT could be converted to a single scale may also have impacted data analysis. While both College Board and ACT publish conversion scales on their websites (ACT, 2008; College Board, 2009), converting all scores to one of the two instruments does not negate the possibility that the two tests measure different aspects of student ability. Results from converted test scores may be problematic, but eliminating students who did not take the ACT from the data set would have decreased sample size and removed many out-of-state students.

As the research was based on pre-existing data, limitations in this study may stem from the lack of control over extraneous variables. The nonrandom sampling method and examination of only one honors program may limit the ability of the study to be generalized to the entire population. In addition, while the overall sample size is quite large, since students self-assigned themselves into groups, equal variances between the cohorts and the application type groups were not realized for certain variables, which affected the statistical analysis of the data. The data for the parental income level and education level was also self-reported, so conclusions regarding socioeconomic status can only be generalized with caution.

Two limitations exist concerning The Honors College data itself. First, The Honors College changed the required first-semester grade point average to maintain eligibility from 3.0 for the 2011 cohort to 3.20 for the 2012 cohort, which may have resulted in greater attrition, but many freshmen who achieve between a 3.0 and a 3.20 the first semester historically became ineligible when they did not achieve the 3.30 required to maintain eligibility after their freshman year. In other words, freshmen may have become ineligible earlier than they would have for the 2012 cohort, but the overall number of ineligible freshmen was likely about the same. Second, it should be noted that The Honors College also invites students to join following the fall semester of their freshman year, based on first semester grade point average (students with a 3.30 or greater are eligible to join, but invitations are only sent to those students with a 3.50 or greater). These students were not included in this study, since they did not submit formal applications. Studying only students who started in the fall semester of their freshman year creates an inaccurate representation of the overall composition of The Honors College, especially since the invitations may result in students from diverse racial or ethnic and socioeconomic backgrounds becoming active participants.

A final limitation impacting the generalizability of the study is that all honors programs use their own unique admission criteria. The information gained from this research can help other

honors programs consider whether implementing an optional essay could benefit their admissions practices, but impacts will vary considerably for each program.

Significance of the Study

The present research contributes to the field of higher education by increasing understanding of Sternberg's (2005; 2010a; 2010b) WICS Model and the augmented theory of successful intelligence (Sternberg, et al., 2004). In particular, this study examines a new population, high-talent students at a public land-grant institution, which differs from previous populations studied using variations of Sternberg's admissions model. Previous studies included the Rainbow Project, administered to all incoming undergraduate students at a private institution (Sternberg, 2010b), the University of Michigan Business School Project, administered to incoming graduate students at a public institution (Sternberg, et al., 2004), and the Kaleidoscope Project, comprised of an optional essay on applications to the private institution Tufts University. Scholars interested in the use of Sternberg's admissions model, and any institutions examining the use of supplementary essays to standardized test scores in admissions, will be interested in the present research.

In addition, the present research contributes to the understanding of honors program admission and retention, which has only been sparsely studied (Moon, 2012). The study increases knowledge regarding the use of minimum scores on the SAT or ACT as admissions criteria for honors programs, and the use of standardized tests for recruiting and retaining high-talent students of diverse backgrounds. The burgeoning numbers of honors programs throughout the United States consider identifying and retaining high talent students a crucial component, and programs may benefit from learning the results of this change to admission policies.

The literature review in chapter two will expand upon the introduction to the study, situating it within previous research and theories.

CHAPTER II

LITERATURE REVIEW

The literature review will begin with a discussion of the history and purpose of honors programs in American universities. It will next discuss the history and use of standardized tests such as the ACT and SAT in college applications, and then describe the two major concerns raised about the use of these standardized tests: validity in predicting first-year GPA, and bias against students from underrepresented ethnic groups and from low socioeconomic status (SES) groups. The literature review will then consider possible implications of the use of standardized tests for honors college admission. Next, the literature review will describe alternative admission initiatives, particularly focusing on Sternberg's (2005) Rainbow Project (2010b), Kaleidoscope Project at Yale University, and Michigan Business School Project (Sternberg, et al., 2004). The literature review will conclude with the theoretical framework for this study.

The History and Purpose of Honors Programs in American Universities

Austin (1986) notes that honors education has "coexist[ed]" with higher education since its inception; honors classrooms today reflect the Socratic discussions of Ancient Greece, the relationships between tutors and students at Oxford University, the seminars of the German research institutions which strongly influenced American institutions, and guild apprenticeships (Austin, 1986, p. 6; Rinn, 2006). American honors programs experienced their first nationwide growth in the 1920s, with Frank Aydelotte (1925) publishing the first report on honors in American higher education in January 1924, at which point 35 institutions were offering honors courses or work; by the publication of Aydelotte's second report in April 1925, this number had

more than doubled to 75 institutions. American honors programs expanded again in the 1950s and 1960s following the Russian launch of Sputnik (Andrews, 2011; Rinn, 2006). Honors programs became a more desirable way to discover and educate the brightest minds, and as a result, new programs developed at many institutions across the country. In addition to the development of new honors programs, the Inter-University Committee on the Superior Student, the precursor to today's National Collegiate Honors Council, began in the late 1950s (Andrews, 2011). The ICSS provided a way for honors programs to convene, learn from each other, establish best practices, and create a kind of standard while still allowing each program to embrace their unique curriculum (Mohler, 2012a).

Interest in and expansion of honors programs in all sizes and types of institutions has recently accelerated. As universities have been forced to compete with each other on a national level, and with the rise of rankings such as the *US News & World Report*, American institutions have encouraged the growth of honors programs, which recruit and retain high-talent students, often raising the prestige of the institution along with the average GPA and SAT or ACT scores of incoming classes (Long, 2002). As of 2002, the majority of honors programs were at public four-year institutions, classified as research or doctoral universities according to the Carnegie system, and listed as "highly competitive," "very competitive," or "competitive" by Barron (Long, 2002, pp. 7-8). It is possible that this demographic has shifted, as honors programs have become popular at community colleges (Selingo, 2002) and little scholarly work has been undertaken in this field.

Institutional benefits of honors programs are rarely discussed, and programs more often focus on the opportunities they provide students: smaller classes, interaction with faculty members, undergraduate research, and oftentimes special scholarships or living communities (Long, 2002). The programs help students who participate in other ways, as well: Cosgrove (2004) used a longitudinal study at three institutions to examine differences between 30 students who persisted in honors until graduation, 82 students who began in honors but did not persist, and

108 students who were eligible for honors but did not participate. He found that students who persisted in honors to graduation had higher grade point averages (3.71 GPA), shorter time to degree, and increased graduation rates (100%), as compared to both students who initially participated but later stopped participation in honors (3.48 GPA for 82% who graduated, 2.76 GPA for 18% who did not graduate), and students who were eligible for honors admission but chose not to participate (3.36 GPA for 76% who graduated, 2.75 GPA for 24% who did not graduate) (Cosgrove, 2004). Moon (2012) surveyed 404 students invited to join an honors program, separated into those who participated and those who did not participate. Responses on a Likert survey indicated honors students were more likely to attend campus art or culture events, work with faculty members on research or independent work, study abroad, and spend more hours preparing for class. She also found that honors participation was a statistically significant predictor of college GPA (Moon, 2012). These results validate some of the numerous but often un-researched claims made by honors programs about benefits to students of participation. As honors programs confer these benefits on only select students, the recruitment and admission of students who will be likely to persist and fulfill requirements is important (McKay, 2009). As part of the effort to recruit capable students, the admission requirements for many honors programs include standardized test scores.

The History and Use of Standardized Test Scores in College Admissions

At the end of the 19th century, colleges gave prospective students their own entrance examinations based on mastery of subject knowledge from typical high school curricula, to determine eligibility for admission (Thelin & Hirschy, 2009). Due to wide variability among the tests and student preparedness, presidents of twelve elite colleges in the Northeast formed the College Board in 1900 to simultaneously standardize the curriculum at the boarding schools their students came from, while giving member colleges the guarantee that students would matriculate with similar capabilities and knowledge (Lemann, 1999). The College Board's essay entrance examinations could not be implemented beyond the Northeast region at the time, so elite colleges

such as Harvard University admitted a few students from outside the region based on high school rank. This changed in 1933, when Harvard began using the Scholastic Aptitude Test as another avenue to find high-talent scholarship students from outside the Northeast (Mohler, 2012b).

The SAT originated from intelligence tests given to nearly two million Army men during WWI, which analyzed aptitude, an inherent trait, rather than achievement, or mastery of subject knowledge (Lemann, 1999). Testing to determine intelligence quotient (IQ) had roots in the field of eugenics; the influx of immigrants to America during this time period reinforced the desire of white elites to protect their unique position at the top of the socioeconomic chain, and intelligence tests seemed to provide scientific support for their perceived superiority. After the war ended, intelligence testers turned to colleges as a new data source with many subjects available for study. The first large group of college students took the SAT in 1926; Harvard's first SAT scholarship class which graduated in 1938 demonstrated the possibility of the test identifying students who would be academically successful, and subsequently other elite colleges and high school students became interested in the SAT. As the number of test-takers increased, the SAT became all multiple-choice, and the earlier College Board essay examinations were phased out during WWII. The American College Test (ACT), first offered in 1959, provided an alternative to the SAT that focused more on subject knowledge as opposed to aptitude (Mohler, 2012b). While no theories validated the use of these standardized tests (Sternberg, et al., 2004), the College Board used years of data to show admissions offices that students' test scores correlated with their first year grade point averages (Lemann, 1999), providing a reason for the continuation of an arbitrarily initiated relationship between colleges and standardized tests.

Today these tests permeate the college admissions landscape, with the SAT more widely used by schools in the western and northeastern United States, and the ACT more widely used by schools in the southern and Midwestern United States (Espenshade, Hale, & Chung, 2005). The tests have changed over the years, and the SAT in particular has emphasized that it no longer tests pure intelligence, but rather subject matter and skills students learn in school (College Board,

2012). However, Frey and Detterman (2004) tested the pre-2005 version of the SAT by correlating the SAT scores of 917 subjects with their scores on the Armed Services Vocational Aptitude Battery (ASVAB), which is generally accepted as a measure of general intelligence, or *g*. They found a strong correlation between the two tests ($r = 0.820$, $p < 0.001$), providing "strong evidence that the SAT is a intelligence test" testing general cognitive ability, not subject knowledge or reasoning ability (p. 374). Regardless of what attribute the tests measure, they are a primary component of college applications in America.

Standardized Test Validity in Predicting First Year Grade Point Average

Beyond equalizing applicants on either general or subject intelligence, standardized test scores are also meant to predict students' first year grades, which could be an indicator of students' likelihood to persist to graduation (Lemann, 1999; Rothstein, 2004). While the College Board initially found the SAT strongly correlated with students' first year grades, later research indicated a correlation around 0.4 between scores and FYGPA, with a correlation of around 0.5 when scores were combined with HSGPA to predict FYGPA (Lemann, 1999). Since the publication of these figures, standardized tests' validity in predicting FYGPA has been the subject of numerous contradictory studies.

Some studies have found the SAT to not be a strong predictor of FYGPA. Geiser and Studley (2002) examined the predictive validity on FYGPA of the SAT I, the traditional aptitude test, versus the SAT II, the less commonly used subject mastery test. By examining standardized regression coefficients for 77,893 students who matriculated in the University of California system from 1996 to 1999, they found that the SAT I predicted 13.3 percent of the variance in FYGPA, while the SAT II predicted 16.0 percent, and HSGPA predicted 15.4 percent. Together, HSGPA and the SAT I predicted 20.8 percent of the variance in FYGPA, compared to 22.2 percent predicted by HSGPA and the SAT II, leading to an overall conclusion that the SAT I should not be accepted on its basis of predicting FYGPA, since other factors can predict FYGPA better. Rothstein (2004) analyzed 18,587 students in California graduating from high school in

1993; he compared their SAT scores with their first year grade point average using goodness-of-fit statistics and found that that the SAT has 20% less predictive power for first year grade point average than the typical assumption of a 0.4 to 0.5 correlation ($R^2 = 0.24$), due to sample selection and statistical error in previous studies which attempt to control for student backgrounds.

Looking beyond predictive validity for FYGPA, Bowen, Chingos, and McPherson (2009) conducted a study concerning personal demographics, academic enrollment, and financial aid of the 1999 cohort at 21 public higher education institutions, including 89,727 first-time freshmen and 19,599 transfer students. Their regression coefficients, based on standard deviations of scores, showed that high school grade point averages alone were a stronger predictor of six-year graduation rates (coefficients ranging from 0.045 to 0.116 depending on school selectivity) than either SAT or ACT scores (coefficients ranging from -0.007 to 0.018 depending on school selectivity). This means an increase of one standard deviation in test scores raises six-year graduation rates less than 2 percent, while an increase of one standard deviation in HSGPA raises graduation rates between 4.5 and 11.6 percent. While Bowen, Chingos, and McPherson (2009) acknowledged that neither the SAT or ACT was designed to predict college graduation, their premise is that the emphasis on first-year retention does not serve students and the public well, because only those students who finish college reap its many benefits.

In contrast to the research advocating for colleges to utilize admissions tools which better predict FYGPA, other research supports the predictive validity of the SAT and ACT. In 2009, Sackett, Kuncel, Arneson, Cooper, and Waters examined statistics for over 150,000 freshmen entering college from 1995-1997, and they determined that SAT scores and FYGPA have a correlation of 0.47. Kobrin, Patterson, Shaw, Mattern, and Barbut (2008) used single and multiple correlational statistics for 151,316 first-time freshmen entering college in 2006, with the new version of the SAT, and found a raw correlation between the SAT and FYGPA of 0.35 (increased to 0.46 when both SAT and HSGPA are correlated with FYGPA). The College Board supported and published the research of Sackett, Kuncel, Arneson, Cooper, and Waters (2009), and the

research of Kobrin, Patterson, Shaw, Mattern, and Barbuti (2008). ACT published the research of Radunzel and Noble (2012), which examined approximately 194,000 first-time freshmen who matriculated at 43 two-year institutions and 61 four-year institutions between 2000 and 2006. Using hierarchical logistic models and regression, they found that both ACT and HSGPA predicted long-term student success, but that only HSGPA was statistically significant when correlated with eventual graduation, while ACT scores had statistical significance when correlated with FYGPA, which then had statistical significance when correlated with eventual graduation. Thus, ACT scores are beneficial in predicting FYGPA, but do not carry the same long-term predictive validity for student success as HSGPA. Rudunzel and Noble (2012) also noted, "ACT does not advocate making college success predictions solely on the basis of a single measure, such as a test score" (p. 11).

The disinterestedness of research regarding standardized test funded by those who profit from their nation-wide use is questionable, but the overall field of SAT and ACT validity testing has numerous examples of contradictory information. The discrepancies between certain findings regarding the tests' predictive power may stem from differences in data gathering and statistical methods (Rothstein, 2004), but the lack of cohesion within the scholarly literature suggests that the ability of the SAT and ACT to predict first year grades, which is the avowed reason for their widespread use in college admissions, has not been definitively proven.

Test Bias Against Minority Students and Students with Low Socioeconomic Status

In addition to discrepancies regarding standardized tests' ability to predict first year grades, scholars have raised questions regarding the apparent bias of tests against students from ethnic and low SES backgrounds (Lemann, 1999; Rothstein, 2004). Differences in scores, especially for black test takers, were first discovered on the IQ tests given to military personnel off of which the SAT was based; the first formal report criticizing the SAT for giving unfair

advantages to primarily white middle and upper class test takers was published in 1948 (Lemann, 1999).

More recently, Rothstein (2004) found from his sample of 18,587 students that SAT scores were more strongly correlated with the characteristics students' schools, including racial composition, percent of students on the free lunch subsidy program, and average parental education level, than were high school grade point averages or college grade point averages. These background characteristics were able to predict, along with HSGPA, 45 percent of the variance in FYGPA, which is similar to what the SAT and HSGPA are able to predict. Statistically controlling for student backgrounds inflated the predictive power of the SAT on first year grade point averages. As his research showed that the SAT was a better measure of a student's background than of their actual preparation for college, Rothstein (2004) summarized his results by wryly suggesting "that admissions offices could admit better-prepared entering classes by giving explicit admissions preferences to high-SES students and to students from high-SES high schools" (p. 315). Boren, Chingos, and McPherson (2009) also found that students' ethnic and socioeconomic statuses were highly correlated with their SAT scores, more strongly than their ethnic and socioeconomic statuses were correlated with HSGPA, to the extent that ethnicity and SES could predict SAT scores with statistical significance. Geiser and Studley (2002) added socioeconomic factors (family income and parental education level) into their regression equations for the University of California data, and found that the predictive validity of the SAT I fell, indicating that the purported relationship between the SAT and FYGPA is explained by socioeconomic factors. Therefore, if institutions control for socioeconomic factors, the SAT I will not provide any predictive information about applicants' FYGPA beyond what HSGPA and the SAT II (which is not as influenced by SES) provide. Scholars often explain such findings by arguing that children from higher socioeconomic backgrounds can afford additional tutoring and test preparation, and are more likely to come from well-funded school districts. Sternberg (2010b) asserts that the reason students from non-white ethnic and lower

socioeconomic backgrounds perform differently on standardized tests is because the SAT and ACT only test one kind of intelligence, analytical, which is most strongly taught and supported by the culture that white middle-to-upper class students are raised in.

In contrast, other researchers have not found that the SAT is strongly correlated with socioeconomic status, or that the predictive validity of the SAT is due to its correlation with SES. Zwick and Green (2007) used secondary data from other analyses (including Geiser and Studley's 2002 UC data) to show that the typical methodology of comparing HSGPA, SAT, FYGPA, and SES is flawed because of the differences between high schools, which are usually aggregated in research looking at a single higher education institution; HSGPA and class rank is more strongly correlated with SES when using within-high schools comparisons, and SAT is less strongly correlated with SES when using within-high school comparisons, as compared to the usual across-high schools methodology. Zwick and Green (2007) dispute the claims that white high SES students perform better on the SAT because the test content conforms to the type of intelligence and worldview those students experience, and that high SES students perform better because of test coaching and expensive test prep courses. Sackett, Kuncel, Arneson, Cooper, and Waters (2009) used correlational statistics and found that while SES is related to SAT ($r = 0.42$), statistically controlling for SES only changes the correlational relationship between FYGPA and SAT from $r = 0.47$ to $r = 0.44$. They used an explanatory model that shows SES influences test results, and test results influence FYGPA, instead of the alternative model assuming SES influences both standardized test results and FYGPA. Sackett, Kuncel, Arnseon, Cooper, and Waters (2009) argue,

> test scores contain meaningful information predictive of academic performance, and the focus shifts to the question of the societal consequences of the fact that being higher in SES confers a meaningful advantage. This may lead some to call for interventions to alleviate the advantage conveyed by high SES. It may also lead some to question test use, but it is important to differentiate between criticizing tests on the ground that they are not valid

> measures of academically relevant skills and criticizing tests on the grounds that one is not comfortable with the social consequences of using a test, despite its being a valid predictor of academic performance. (p. 2)

In their view, the SAT-SES correlation is not a reason to discontinue use of the test, because the SAT-FYGPA correlation still shows that the test's predictive validity is strong. Zwick and Green (2007) echo this sentiment with the observation that "until we have a socioeconomically equitable society, we will not have a socioeconomically neutral test" (p. 43).

Despite the conflicting evidence regarding the bias of the SAT and ACT against low SES and ethnic students, they are still widely used by universities for admissions purposes. This raises concerns regarding how institutions can promote diversity among student bodies, especially as legal cases have prohibited universities from using affirmative action to admit students from different ethnic backgrounds over non-minority students with better standardized test scores (Gratz v. Bollinger, 2003; Hopwood v. Texas, 1996; Lemann, 1999; Regents of the University of California v. Bakke, 1978).

Admission Criteria and Implications for Honors Programs

Like the institutions within which they reside, honors programs struggle to find equitable tools to admit students who are qualified and likely to succeed. Admission criteria for honors programs vary across the nation, but as of 2002, around 50% of honors programs utilized the SAT or ACT to consider students for admission (Long, 2002). Many honors programs use minimum required scores on the SAT or ACT test as part of their admissions criteria (McKay, 2009). Despite this widespread practice, the ability of the SAT or ACT to predict success in honors programs has not been widely studied. McKay (2009), in studying 1,017 honors students using logit regressions, found that high school grade point average was a better predictor for honors program completion than SAT scores. Moon (2012) found that honors students had statistically significant higher results than non-honors students on HSGPA, ACT scores, mother's education, father's education, and cumulative college GPA, and that in addition to honors participation,

mother's education, HSGPA, and ACT score, were all statistically significant predictors of college GPA for both honors and non-honors students. Khe (2007), in contrast, analyzed five years of matriculating honors students in an honors program that admits based on minimum HSGPA and standardized test score, and found "no consistent pattern" between high school achievement and college achievement, as shown by college GPA (p. 79). Smaller honors programs are able to use individual interviews to look for talents not usually expressed by HSGPA or standardized test scores, such as intrinsic motivation or intellectual curiosity; Cosgrove (2004) suggests that finding a way to consider students' motivation and goals as part of admissions may improve the chances of honors programs selecting students who will be more likely to persist and benefit from the unique experience.

Of greater concern than the ability of the minimum standardized test scores to predict success in honors, however, is the potential bias of the test scores outlined above (Lemann, 1999). The use of minimum scores increases the likelihood that students from certain backgrounds, notably ethnic minorities and low socioeconomic status, are not eligible to participate, and thus cannot reap the benefits these programs provide through increased individualized attention. Thus, honors practitioners question the use of standardized tests as a way to admit qualified students, while also highlighting the importance of increasing honors student diversity. The National Collegiate Honors Council's "Beginning in Honors: A Handbook" (Schuman, 2006) affirms:

> If an honors program offers itself as the institution's best, it needs to pay careful attention to the racial, ethnic, class, and geographical demographics of its student body and its faculty and staff. Today's colleges and universities, at their best, offer students opportunities to learn, informally as well as officially, with others from a wide variety of backgrounds. A persuasive case can be made that honors programs should be more diverse and should have more students of color, more international students, show a greater geographical mix, and attract students with a wider variety of personal styles and preferences than the institutions that house them. This demographic has, in fact,

> sometimes been the case. At other times, unfortunately, honors programs have had difficulty sustaining diversity in their student population. (p. 24)

There is little to no research on admissions methods to increase diversity in honors, beyond honors educators asserting that the opportunities provided by honors programs should be open to a wide variety of talented students in order to create the best learning environment and the strongest scholars and future leaders possible (Cosgrove, 2004; McKay, 2009). McKay (2009) explicitly argues against using standardized testing as a minimum requirement for acceptance to honors programs, because it "may reduce the diversity of a program and falsely exclude qualified demographics," especially by "exclud[ing] more diverse types of thinkers who do not do well on standardized tests but who would otherwise achieve great success" (pp. 83-84). For larger honors programs whose admissions policies cannot include expensive and time-consuming individual interviews, finding alternative or augmentative admissions processes beyond HSGPA and standardized tests, to identify students from diverse backgrounds who will excel academically and persist in honors, is an exigent challenge.

Alternative Admissions Initiatives

The conflicting research regarding the predictive validity of standardized tests, and their potential bias against minorities and students from low socioeconomic backgrounds, have led some educators to seek alternative or augmentative admissions policies as an equitable solution. In 2001, the University of California system stopped requiring the SAT or ACT for student admission (Atkinson & Geiser, 2009). In 2008, the National Association for College Admission Counseling released the Report of the Commission on the Use of Standardized Tests in Undergraduate Admission,

> encourag[ing] institutions to consider dropping the admission test requirements *if it is determined that the predictive utility of the test or the admission policies of the institution (such as open access) support that decision and if the institution believes that*

> *standardized test results would not be necessary for other reasons such as course placement, advising, or research* [emphasis original]. (p. 7)

Institutions must consider the use of standardized scores given their own mission; liberal arts colleges and other elite institutions do not rely on standardized test scores as much as public schools with larger enrollment (Zwick, 2007), and over 280 not-for-profit four year institutions do not require students to submit ACT or SAT scores for admission (National Association for College Admission Counseling, 2008). Some institutions have considered alternative admissions systems, such as admitting by class rank or using a lottery system, but using external findings from other sources, Zwick (2007) argued that these options were no more effective in garnering a diverse group of students than traditional admissions methods utilizing grade point averages and standardized test scores.

The Rainbow Project

Sternberg (2010b) used the WICS (Wisdom, Intelligence, Creativity, Synthesized) model to generate a series of empirical studies to identify students with other kinds of intelligence beyond analytical, as well as students from diverse ethnic and socioeconomic backgrounds. The Rainbow Project in 2001 tested 777 first-year college students at 13 different higher education institutions for analytical, creative, and practical skills using a variety of tests designed by the researchers to augment the SAT. Using hierarchical regression, the researchers found that the students' scores on the analytical, practical, and creative skill tests doubled the predictive power of the SAT test for FYGPA, that Rainbow Project scores could account for 20 percent of the variance in FYGPA without SAT or HSGPA included in regression calculations, and that the Rainbow Project tests reduced the deviation from the mean group differences among students from non-white ethnic backgrounds (Sternberg, 2010a; Sternberg, The Rainbow Project Collaborators, and the University of Michigan Business School Project Collaborators, 2004).

University of Michigan Business School Project

The University of Michigan Business School Project explored the use of Sternberg's theory of successful intelligence to augment the GMAT (Graduate Management Admission Test, administered by the Graduate Management Admission Council, with similar question types and format as other standardized tests) (Sternberg, The Rainbow Project Collaborators, and the University of Michigan Business School Project Collaborators, 2004). The researchers developed their own case study and situational-judgment questions involving analytical, practical, and creative intelligence, which they administered to 422 first-year Master of Business Administration students in 1999. The correlational statistics showed that students who performed well on the newly developed questions, which did not correlate with GMAT scores and thus tested for entirely different abilities, had higher first year GPAs, were more involved in student organizations and student leadership positions, and eventually received more job offers upon graduation. In addition, the newly developed questions reduced gender and ethnic disparities; for instance, blacks scored 0.14 standard deviations lower than whites on the situational-judgment questions, as compared to 1.24 standard deviations lower on the GMAT (Sternberg, The Rainbow Project Collaborators, and the University of Michigan Business School Project Collaborators, 2004).

The Kaleidoscope Project

Starting in 2006, Sternberg (2010b) initiated the Kaleidoscope Project at Tufts University to implement pieces of the Rainbow Project and his newly developed WICS theory into college admissions, using optional essays designed to exhibit students' creativity, wisdom, practicality, or analytical skills. The optional essays supplemented but did not replace standardized test scores. Data for the 2011 graduates of Tufts showed the ethnic diversity of those applicants accepted to Tufts increased (black student acceptance increased 30% in the first year of Kaleidoscope, and Hispanic student acceptance increased 15% in the first year of Kaleidoscope), and students who wrote the optional essays were more likely to have higher FYGPAs and be involved on campus as

leaders. Scores on the Kaleidoscope did not correlate with SAT scores, proving that the project looks for entirely different abilities of applicants, but Kaleidoscope scores did correlate with extracurricular involvement and leadership. Kaleidoscope results did not show the same doubling of the predictive ability of the SAT as the Rainbow Project (Sternberg, 2010b).

Theoretical Framework

Sternberg's (2005) WICS model provides the main theoretical framework for this research study. The WICS model was chosen because it best fits the research purpose and questions. Other theories that could have been chosen include Howard Gardner's (2011) theory of multiple intelligences, which posits that humans have nine distinct kinds of intelligences: bodily-kinesthetic, existential, interpersonal, intrapersonal, linguistic, logical-mathematical, musical, naturalist, and spatial intelligences, all of which "meet certain biological and psychological specifications" (Gardner, 2011, p. 66). Gardner contends that educators and standardized test developers focus mainly on linguistic and logical-mathematical intelligences, and the educational system should recognize that individuals have distinct strengths based on their genetic and cultural backgrounds. While Gardner's multiple intelligences provide more breadth than Sternberg's (2005) four kinds of intelligence (creative, analytical, practical, and wisdom), Sternberg's theory more readily lends itself to assessment and testing; Sternberg has developed many assessment tools to explore the WICS model, including its use in college admissions, whereas Gardner (2011) has not tested or revised the theory of multiple intelligences much in the thirty years since its inception. The essay questions in the current study were developed by The Honors College with Sternberg's (2005) theory in mind, to reveal creative, analytical, practical, and wisdom-based intelligences in prospective students. Bringing in an unrelated theory with superfluous intelligences would have lent less structure and insight into the current research study than Sternberg's WICS model.

Sternberg's (2005) WICS model postulates that leaders need different kinds of intelligence and skills in order to be successful. Leaders need creativity in order to generate novel

ideas. They need both analytical and practical intelligence, analytical to understand and process information, and practical to carry out ideas. Leaders also need wisdom in order to use the previous skills to benefit humanity, instead of using their leadership position for immoral or corrupt purposes. Sternberg maintains that contemporary educational systems only recognize and reward analytical intelligence, which results in the education and promotion of leaders who only have one strong skill set, while others who might have a better balance of leadership traits are not recognized and placed in an optimal position for society to benefit from their potential. This criticism extends to standardized testing, which Sternberg does not aim to completely replace, but to augment with admissions measures that seek out students and potential leaders with other kinds of intelligence. Sternberg derived the WICS model from other leadership theories, especially transformational leadership (Northouse, 2010). Transformational leadership differs from WICS in that it focuses mainly on how charismatic leaders help their followers achieve more than the followers thought possible; the exact requirements for great leaders are not as succinctly described in transformational leadership. The WICS model provides a more concrete way to understand how to educate people, building on their inherent strengths, to eventually contribute positively to humanity. The WICS model has evolved out of Sternberg's previous work, including the triarchic model to identify and assess gifted children (Sternberg, Ferrari, Clinkenbeard, & Grigorenko, 1996), and the augmented theory of successful intelligence (Sternberg, et al., 2004), which did not discuss wisdom.

In addition to providing the theoretical framework for this research by outlining why it is important for universities to recruit and admit students with different skill sets, Sternberg's work also provided the framework for the methodology, and the selection of the variables studied. This study in essence replicated Sternberg's Kaleidoscope Project from Tufts University, but with a different population. The research questions for this study were based on findings from Sternberg's previous work. The optional essays used by The Honors College were derived from

Sternberg's (2005; 2010b) work and were intended to increase the diversity of students by supplementing standardized test scores.

Summary of Literature Review

Honors programs, started by many institutions to attract high-talent students (Long, 2002), oftentimes use minimum required scores on the SAT and ACT as part of their admissions process (McKay, 2009). The SAT and ACT tests rose to prominence during the 20th century for college admissions in America (Lemann, 1999), though their ability to accurately predict student grades and college retention has been questioned (Rothstein, 2004), and the tests are widely believed to be biased against ethnic minorities and students from low socioeconomic backgrounds (Lemann, 1999; Rothstein, 2004). Alternate methods of admission have been suggested, including lottery admission, percentile admission, and Sternberg's (2010a; 2010b) optional essays intended to highlight students' creativity, practically, and wisdom in addition to their analytical skills.

CHAPTER III

METHODOLOGY

The purpose of the present study was to examine potential differences in background characteristics and academic achievement of honors students who submitted different types of applications. The methodology section will first state the research questions, then will describe the research design, participants in the study, procedures, and will conclude with methods of analysis.

Research Questions

The present study sought to answer the following research questions:

1. Do students who submitted augmentative applications differ in race or ethnicity, socioeconomic status, or academic achievement from students who submitted traditional applications?
2. Are students who submitted augmentative applications retained differently than students who submitted traditional applications?

Research Design and Instrumentation

This study does not fit into one particular subgroup of quantitative research; it blends causal-comparative research and correlational research, in the sense that it seeks to explain relationships or patterns using secondary analysis of pre-existing data, but without determining causation (Gay, Mills, & Airasian, 2011). The design allowed the researcher to determine what, if any, differences exist between two matriculating classes in The Honors College. Any statistically

significant relationships discovered require further research using experimental designs if possible, to determine causation.

The institution at which the study took place gathered the pre-existing data for this research. The creation of the data for most of the variables did not involve a specific instrument beyond institutional forms for students to fill out or submit (HSGPA, race or ethnicity, gender, reported parental income, FYGPA, and fall-to-fall retention). SAT or ACT scores resulted from standardized tests given to the students during their high school careers. While the creators of the tests label them achievement tests, researchers have described them as aptitude tests examining only one quality, intelligence (Frey & Detterman, 2004; Lemann, 1999; Rothstein, 2004).

Many scholars have studied the reliability and validity of standardized tests without conclusive findings for either continuing or discontinuing their use. The Honors College derived the optional application essays from Sternberg's (2005; 2010b) Rainbow Project and Project Kaleidoscope, whose reliability and validity have not been extensively researched (Sternberg & The Rainbow Project Collaborators, 2006). Unlike Project Kaleidoscope, The Honors College did not assign numerical scores to essay applications, so statistical analysis methods will differ.

Definitions of variables and key terms

The operational definitions of the variables studied and other key terms, within the context of The Honors College, are as follows:

Variable or Key Term	Operational Definition
Fall-to-fall retention	Students are retained when they enroll at an institution and remain enrolled in subsequent semesters (Tinto, 1993). Fall-to-fall retention looks at whether students who enroll in their first fall semester are still enrolled one year later in their second fall semester. In this study, students who were retained were enrolled and active in the fall semester of both their freshman and sophomore year. Active status in The Honors College requires freshmen and sophomores to enroll in 6 honors hours each semester; students enrolled in less were not included in the retention count.
First-year grade point average (FYGPA)	A students' grade point average is determined by grades earned in courses; per credit hour, students receive 4 points for an A, 3 points for a B, 2 points for a C, 1 point for a D, and 0 points for an F. Total points are divided by credit hours earned to determine the grade point average. Students' grade point averages for all college coursework (excluding pass/fall credit such as AP tests) were collected at the end of the spring semester of their freshman year.

Race or ethnicity	Race or ethnicity can refer to biological or sociological factors such as physical appearance, genetics, cultural group, or ancestral heritage, that categorize groups of people in the world. Students reported their racial and ethnic information during their application process. Students could choose from Alaskan native/American Indian, Asian, black/African American, Hispanic/Latino, white, multi-racial, or they could decline to identify. If a student chose Hispanic as their ethnicity, their racial choice was also reported as Hispanic.
Gender	Students reported their gender information during their application process. They could choose between male and female.
Reported parental income	The institution used the Department of Education FISAP (Fiscal Operations Report and Application to Participate), which includes information regarding family income. Dependent students report total income and student's total income, while independent students report only their own total income. Since this data was self-reported on financial-aid related materials, caution must be used regarding generalizations for socioeconomic status.
Standardized test scores	Standardized tests are nationwide tests given to all students under the same circumstances, in order to provide a fair method of comparing students from different backgrounds on the same measure. Most colleges require a standardized test score as part of students' applications. The two most common standardized tests are the ACT and SAT. Both tests contain multiple choice questions designed to determine students' mastery of high school subjects such as mathematics, critical thinking, reading, writing, and science. All scores for the present study were converted to the ACT scale.
High school grade point averages (HSGPA):	A students' grade point average is determined by grades earned in courses; typically per credit hour, students receive 4 points for an A, 3 points for a B, 2 points for a C, 1 point for a D, and 0 points for an F. Total points are divided by credit hours earned to determine the grade point average. Some high schools award more points for Advanced Placement or International Baccalaureate grades, so that a student earning an A may receive 5 points per credit instead of 4. Other high schools do not follow a 4.0 scale. All applicants' HSGPAs were converted to a 4.0 scale for consistency.
Traditional Application	An application to The Honors College that included only standardized test scores and high school grade point average.
Augmentative Application	An application to The Honors College that included the two elements of traditional applications, plus an optional essay.

Research Context

The location of the study was a large, four-year, land grant research university classified as selective (Carnegie Foundation for the Advancement of Teaching, 2010). The university is in the Midwest region of the United States. There are approximately 23,000 enrolled students, of which approximately 19,000 are undergraduates (Present Student Body, 2012). 48 percent of undergraduates are female, and 52 percent are male. 73 percent of undergraduates are state residents, while 24 percent are residents of other states, and 3 percent are international students. 76 percent of undergraduates are not ethnic minorities, and 24 percent are ethnic minorities. As a percentage of total undergraduate population, African Americans comprise 4.7 percent, Native

Americans 6.2 percent, Hispanics 4.3 percent, Asians 1.5 percent, and 6.6 percent identified as Multiracial. Students may choose from over 100 undergraduate majors in six colleges, including agricultural sciences, arts and sciences, business, education, engineering, and human sciences. The institution is located in a town of approximately 45,000 residents. Though within an hour of two larger cities, the majority of students live on campus or in the surrounding town.

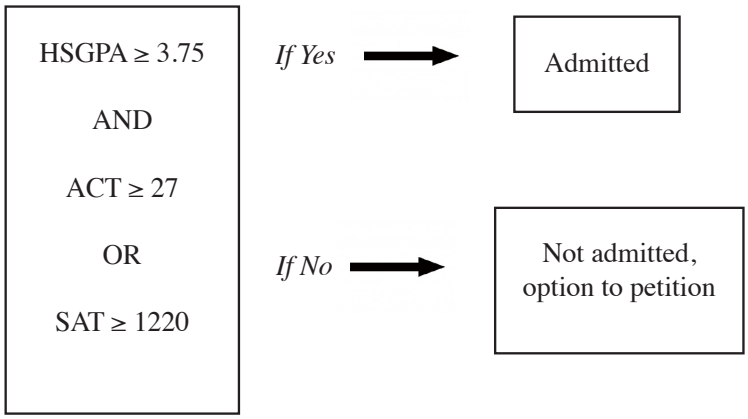
Participants

The population for this research encompassed all honors students at the institution studied. The Honors College typically includes 5-6 percent of the entire undergraduate body. This population did not include those who applied to participate, but either did not gain admission or chose not to participate in The Honors College.

Students submitted a separate application (with no additional application fee) to The Honors College. The Honors College could not admit students until the university admissions office had accepted them. Honors applications had no influence on applications to the university. Prior to the 2012 – 2013 freshman class, students could apply to The Honors College using traditional applications (illustrated by Figure 3.1 below). Students with a 3.75 HSGPA, weighted or un-weighted, and either an SAT of 1220 (critical reading and math only) or higher or a composite ACT of 27 or higher would be admitted. Students short of these requirements could submit a petition for further consideration, which included academic transcripts, a letter of recommendation, and a personal statement. Faculty in students' intended major colleges made decisions regarding petition applications.

Starting for the 2012 – 2013 incoming freshmen, an optional essay based on Sternberg's (2005; 2010b) WICS model could augment the traditional applications. The five optional essay questions for 2012-2013 students are listed in Appendix 1. Students could respond to any of the five questions with an essay of 250-400 words. The Honors College automatically accepted students who met the traditional application criteria, whether they submitted an augmentative

Figure 3.1: Traditional Application Process

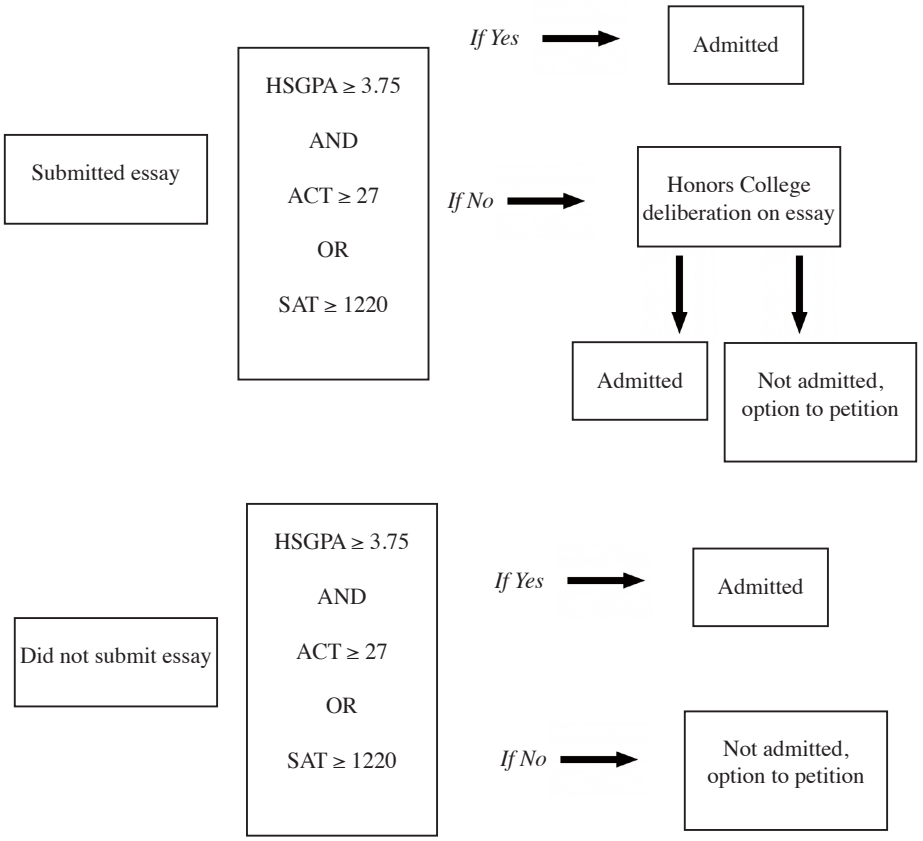


essay or not. The Honors College staff deliberated over the essays written by students who did not qualify for traditional admission. Students whose essays showed intelligence in the areas of wisdom, creativity, practicality, or analysis beyond that demonstrated by the traditional application scores could be admitted using the augmentative applications, as outlined in Figure 3.2 below. This process mostly replaced petitions, though students could still petition for admission to The Honors College for the 2012-2013 year if their applications were denied.

The researcher selected the sample for this study using convenience sampling; the sample included all of the individuals within two consecutive incoming classes of The Honors College. The nonrandom sample selection may not accurately represent the population as a whole (Gay, Mills, & Airasian, 2011).

The sample included two subgroups: the 2011-2012 entering class of The Honors College, and the 2012-2013 entering class. These subgroups only included freshmen enrolling in their first semester of college; the sample did not include transfer students or adult students with prior college credit. The sample included any student active in their first semester (enrolled in six honors hours) for the purpose of tracking retention, regardless of age, gender, ethnicity, or major. The 2011-2012 incoming class had 456 active first-semester freshmen; the 2012-2013 class had 530 active first-semester freshmen, leading to a total sample of 986 students.

Figure 3.2 Augmentative Application Process



Procedures

The researcher collected data from two separate databases at the institution. The institution collects data for SAT or ACT scores, race or ethnicity, gender, reported parental income, FYGPA, and HSGPA in the campus-wide database Student Information Systems (SIS). The Honors College keeps a separate Filemaker database which houses information on whether students submitted optional essays, whether students met normal admission requirements, and active status for each semester. To compile data from these two sources, first an Honors College employee other than the researcher exported the necessary data from Filemaker into an Excel spreadsheet, along with CWIDs (Campus Wide Identification numbers), and assigned randomized numbers found online using the website random.org. The employee then sent this spreadsheet to the Institutional Research and Information Management (IRIM) office on campus. IRIM matched

the information from Filemaker with the information from SIS, de-identified the data by deleting all CWIDs, and sent the spreadsheet back to the researcher. The researcher then had access to all of the necessary data, completely de-identified, and therefore did not need to obtain consent from students to undertake the present research project. Before beginning analysis, the researcher converted SAT and ACT scores to the same scale using concordance tables available from both College Board and ACT online. Since ACT is more prevalent for in-state students, SAT scores were converted to the ACT scale.

The researcher began analyzing data for the 2011 entering class in summer 2013, and began analyzing the data for the 2012 entering class after the third week of the fall 2013 semester, at which point The Honors College had determined active status, and thus fall-to-fall retention, for the entering class of fall 2012. This necessitated two separate spreadsheets sent to IRIM for data compilation, with distinct randomized numbers to eliminate confusion.

Analysis of Data

Two separate comparisons were made. The first comparison looked at two cohorts of honors students: those admitted using traditional applications in the 2011-2012 academic year, and those admitted for the following year when augmentative applications based on Sternberg's (2005; 2010a; 2010b) research became an option for applicants. The second comparison looked only at students admitted in the 2012-2013 academic year when augmentative essays were added, but divided the students into three groups: those who submitted a traditional application without an essay, those who qualified for submitting a traditional application but chose to submit an augmentative application with an essay, and those who did not qualify for the traditional application and were admitted based on their augmentative application. The groups in both comparisons were analyzed on multiple dependent variables, including fall-to-fall retention, first-year cumulative grade point average, race or ethnicity, gender, reported parental income, standardized test scores, and high school grade point average.

The researcher analyzed the data using statistical methods in SPSS. For the interval and ratio dependent variables (FYGPA, SES, SAT/ACT scores, and HSGPA), the researcher conducted a factorial ANOVA test, with two independent variables (matriculation year, and application type). For the nominal dependent variables (retention, diversity, and gender), the researcher conducted separate Pearson's Chi-Square tests, for each of the independent variables. Statistical significance required meeting the critical *p* value of 0.05, indicating that the probability of observing the reported results by chance was less than 5 percent. For statistically significant ANOVA results, the researcher conducted Tukey's post hoc test to determine where differences existed for interval and ratio dependent variables.

CHAPTER IV

FINDINGS

The purpose of the present study was to examine potential differences in background characteristics and academic achievement of honors students who submitted different types of applications. Students who applied to the honors college using traditional applications (high school grade point average and standardized test scores) were compared to students who applied to the honors college using augmentative applications (adding an optional essay component to the traditional application). Research questions for the study included:

1. Do students who submitted augmentative applications differ in race or ethnicity, socioeconomic status, or academic achievement from students who submitted traditional applications?
2. Are students who submitted augmentative applications retained differently than students who submitted traditional applications?

It was hypothesized that students who submitted augmentative applications would differ from students who submitted traditional applications, and that augmentative applicants would be retained differently than students who submitted traditional applications. The results supported the first part of the hypothesis for two variables: the race or ethnicity of students in the 2011 cohort differed from students in the 2012 cohort, and students who submitted traditional/augmentative applications had lower ACT scores than students who submitted traditional applications. The results supported the second part of the hypothesis as well; a lower percentage of students who submitted petition and augmentative applications were retained into

their second fall compared to students who submitted traditional and traditional/augmentative applications, and no difference in retention was found between students who submitted traditional and traditional/augmentative applications. In addition, students who submitted augmentative applications had lower first year grade point averages than all other students. Cohort year was not significantly related to either retention or first year grade point average.

The findings section will first summarize the results of hypothesis testing, and then present the statistical analyses for each of the two research questions, grouped by independent variable and then by dependent variable.

Differences by Cohort Year

To test the hypothesis regarding differences between the two matriculating classes, Pearson's Chi-Square and factorial ANOVA tests were run with cohort year as the independent variable. The following section details the test results.

Gender

The numbers of females and males matriculating in each cohort did not differ in a meaningful way. Figures 4.1 below illustrates that the proportions of females and males stayed about the same, with 53.51 percent females in 2011 and 53.58 percent females in 2012. The results of a Pearson's Chi-Square test with gender as the dependent variable were not statistically

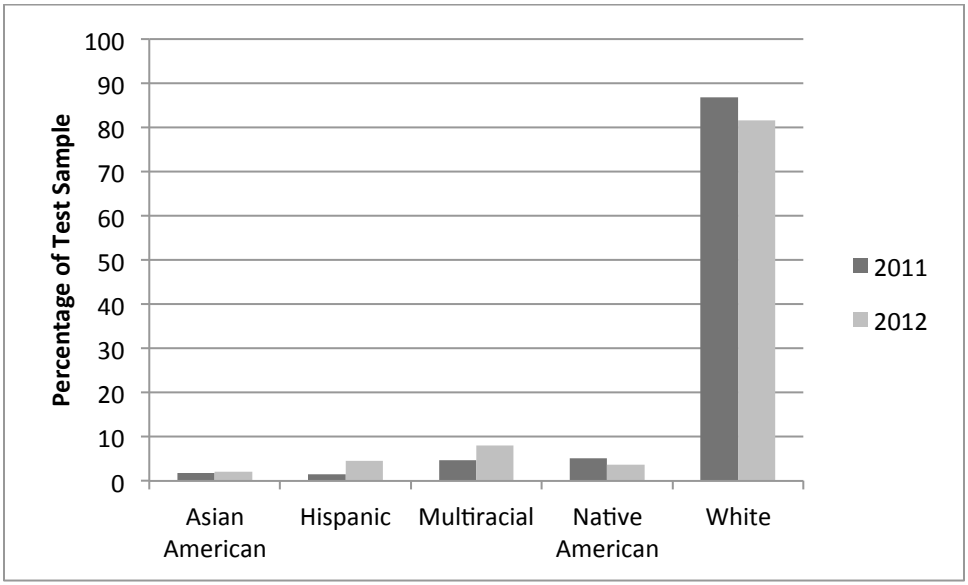
Figure 4.1 Gender Distribution by Cohort Year

significant, X^2 (1, N=986) = 0.001, p = 0.98. The results indicate that adding an optional essay did not lead to a change in the proportion of men and women in The Honors College.

Race and Ethnicity

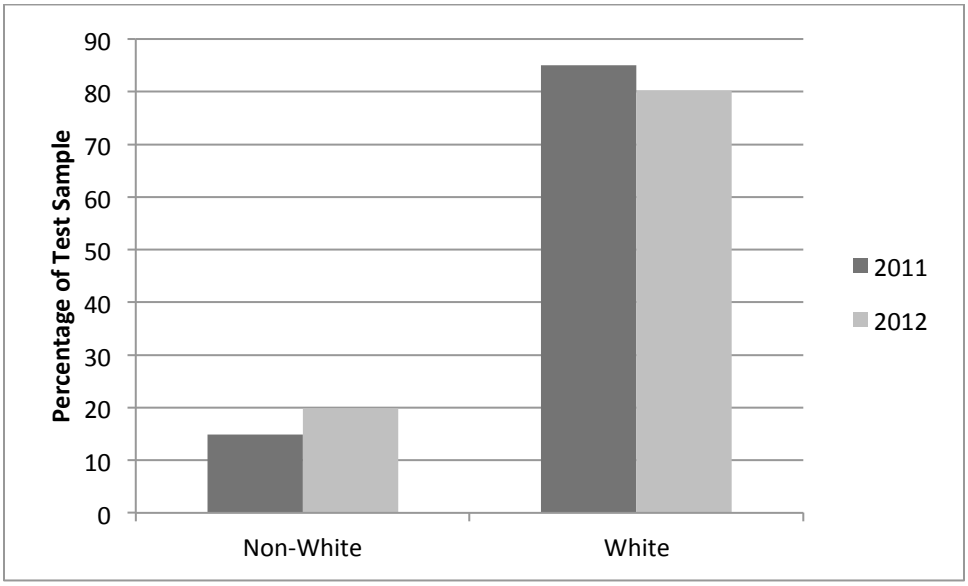
The proportion of matriculating students who self-identified as Asian American, Hispanic, Multiracial, Native American, and White differed significantly from the 2011 cohort to the 2012 cohort. From 2011 to 2012, the percentage of Asian American students increased from 1.79 to 2.11; the percentage of Hispanic students increased from 1.57 to 4.61; the percentage of Multiracial students increased from 4.70 to 8.06; the percentage of Native American students decreased from 5.14 to 3.65; and the percentage of White students decreased from 86.80 to 81.57 (see figure 4.2 below). The overall proportion of non-white ethnicities in the 2011 cohort

Figure 4.2 Race and Ethnicity Differences by Cohort Year



differed significantly from the 2012 cohort. Figure 4.3 below illustrates that from 2011 to 2012, the percentage of non-white students increased from 14.91 to 19.81, and the percentage of white students decreased from 85.09 to 80.19.

Figure 4.3 Non-White and White Ethnicity Differences by Cohort Year



Due to low responses for African Americans, international students, and unknown, these categories were removed for the first Pearson's Chi-Square test (see Table 4.1 for exact figures for the distribution of ethnicities). The first test with Asian American, Hispanic, Multiracial, Native American, and White ethnicities as the dependent variable, and cohort year as the independent variable yielded statistically significant results, X^2 (4, N=968) = 13.28, p = 0.01. A Cramer's V effect size of 0.01 with 4 degrees of freedom indicates a small effect size for the relationship between matriculation year and ethnicity.

Table 4.1 Ethnicity Distribution

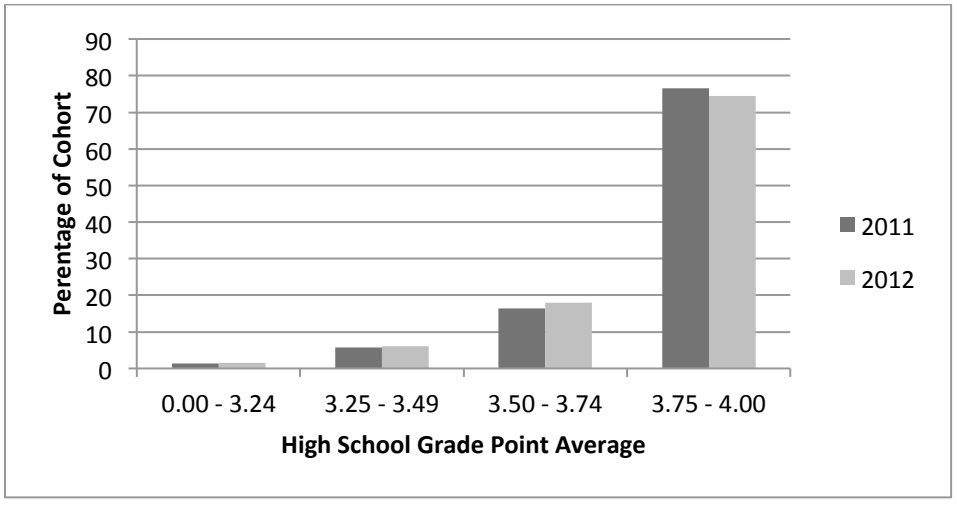
	2011 Cohort	2012 Cohort
African American	3	6
Asian American	8	11
Hispanic	7	24
International	4	0
Multiracial	21	42
Native American	23	19
Unknown	2	3
White	388	425

A second Pearson's Chi-Square test conducted with white and non-white ethnicities as the dependent variable also yielded statistically significant results, X^2 (1, N=986) = 4.07, p = 0.04. A Cramer's V effect size of 0.04 with 1 degree of freedom indicates a small effect size, meaning the relationship between matriculation year and ethnicity was stronger when all non-white ethnicities were considered in aggregate, but still not very large. The results indicate that in the first year students could choose to submit optional essays, the proportion of White and Native American students significantly decreased, while the proportion of Asian American, Hispanic, and Multiracial students significantly increased. The overall proportion of non-white students increased among the 2012 cohort.

High School Grade Point Average

The students who matriculated in 2011 (M = 3.85; SD = 0.20) and 2012 (M = 3.82; SD = 0.31) had similar HSGPAs (see Figure 4.4 below). Though the percentage of students with

Figure 4.4 High School Grade Point Average Ranges by Cohort Year



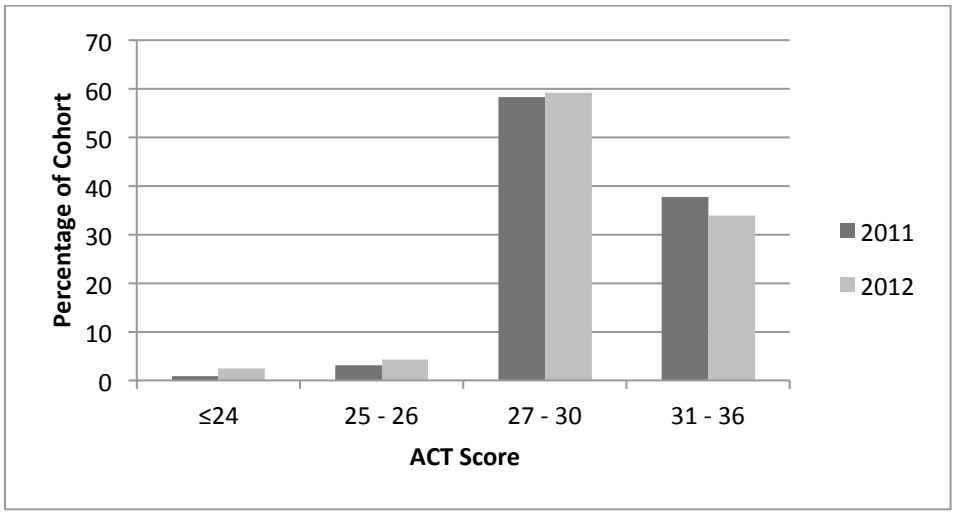
HSGPAs lower than 3.75 increased slightly, the overall percentage of students with the minimum 3.75 HSGPA or above stayed about the same, 76.54 percent in 2011 and 74.53 percent in 2012. A factorial ANOVA with HSGPA as the dependent variable and cohort year, application type, and

retention as independent variables found that the cohort year did not have a statistically significant relationship with high school grade point average, F (1, 986) = 1.66, p = 0.20. The results indicate that the high school academic achievement of students entering The Honors College did not change in the first year students had the option to submit essays with their applications.

ACT Scores

The 456 students who matriculated in 2011 (M = 29.99; SD = 2.40) and the 530 students who matriculated in 2012 (M = 29.58; SD = 2.54) had similar ACT scores (see Figure 4.5 below).

Figure 4.5 ACT Score Ranges by Cohort Year

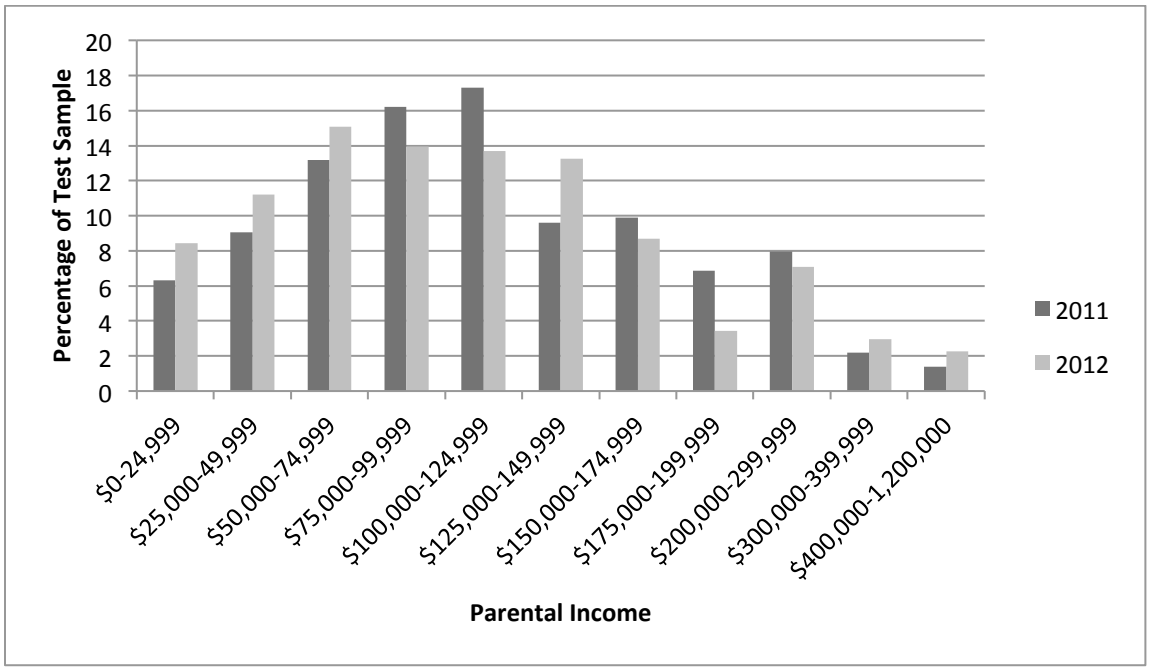


In 2011, 96.05 of students had the minimum 27 ACT score; in 2012, this percentage decreased to 93.21 percent. A factorial ANOVA conducted with ACT scores as the dependent variable and cohort year, application type, and retention as the independent variables found that ACT scores were not related to cohort year with statistical significance, F (1, 986) = 1.47, p = 0.23. These results indicate that the ACT scores of the incoming honors students did not change when students had the option to supplement their applications with optional essays.

Parental Income

The 364 students who matriculated in 2011 (M = \$124,525.05; SD = \$97,869.14) and the 438 students who matriculated in 2012 (M = \$119,821.03; SD = \$100,144.91) had similar parental incomes (see figure 4.6 below). The percentage of 2012 students whose parents made

Figure 4.6 Parental Income by Cohort Year



\$74,999 or less was higher than the percentage of 2011 students, but this pattern is not consistent for higher salary amounts. A factorial ANOVA conducted with reported parental income as the dependent variable and cohort year, application type, and retention as the independent variables found that parental income was not related to cohort year with statistical significance, F (1, 802) = 0.19, p = 0.66. The results indicate that the socioeconomic status of incoming honors students, as represented by parental income, did not change when students had the option to submit an essay as part of their application.

Summary

The students in the two cohorts did not differ on variables that described their backgrounds prior to their matriculation in college, such as gender, academic achievement in high school, and socioeconomic status. The findings did not support the hypothesis that students in the two cohorts would differ, for most variables. Adding an optional essay to Honors College applications did not lead to many significant differences between the last incoming class that did not have the option to write the essays and the first incoming class that had the option to write the essays. It is imperative to note that the research design did not allow the researcher to determine causality, only to examine where variables were related to each other. While significant findings could have indicated that the essays were related to a change or difference in the kinds of students admitted to The Honors College, for the most part the two cohorts were the same.

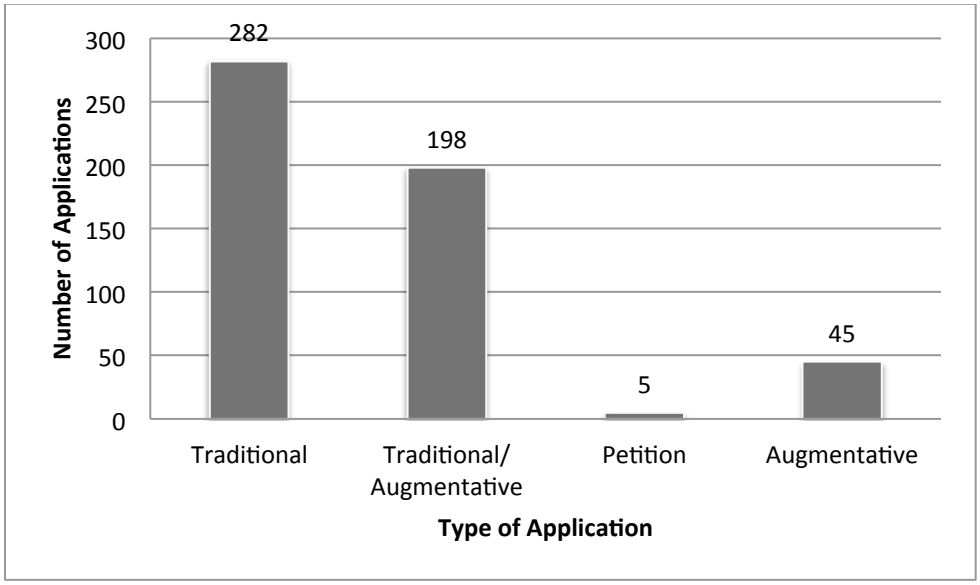
The only variable on which the 2011 and 2012 cohorts differed was race or ethnicity. The increase in non-white minority students among the 2012 class could potentially be related to the essays, as previous research using optional essays based on Sternberg's WICS model led to increases in non-white students (Sternberg, 2010a; Sternberg et al., 2004). Other factors could also have attributed to the increase in non-white ethnicities among the 2012 cohort. For instance, the institution studied had the largest incoming class in its entire history in 2012. University-wide recruitment of incoming freshmen, available tuition and financial aid, national trends in the economy, popularity of the institution due to athletic teams, or other unknown factors could have impacted the increase in non-white ethnicities in The Honors College in 2012.

Differences by Application Type

To test the hypothesis regarding differences between students submitting different kinds of applications, Pearson's Chi-Square and factorial ANOVA tests were run with application type as the independent variable. The following section details the test results, grouped by dependent variable. All of the data in this section concerns only the 2012 cohort, who had the option to choose augmentative applications.

Figure 4.7 illustrates the types of applications submitted by the 530 students in the 2012 cohort. Traditional applicants met the minimum HSGPA and ACT admissions requirements and

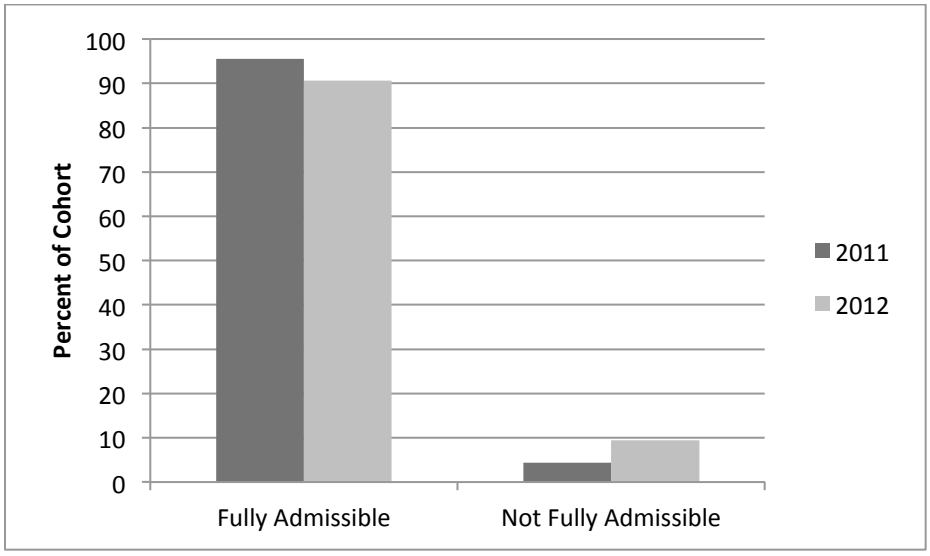
Figure 4.7 Application Types Submitted by 2012 Cohort



did not submit an optional essay; traditional/augmentative applicants met the minimum admissions requirements and did submit an optional essay; and augmentative applicants did not meet the minimum admissions requirements and did submit an essay. Of the 530 students admitted for the 2012 cohort, 243 students (45.85 percent) submitted essays; 81.48 percent of essay applicants were already admissible based on their HSGPA and ACT score.

Students whose initial application to The Honors College was not accepted still had the option to submit a petition, which included high school transcripts, a letter of recommendation, and a cover letter. The augmentative applications mostly replaced petitions for the 2012 cohort. Figure 4.8 below illustrates that the percentage of students admitted who did not meet the 3.75 HSGPA and 27 ACT minimum requirements increased from 4.39 percent in 2011 to 9.43 percent

Figure 4.8 Admissibility by Cohort Year



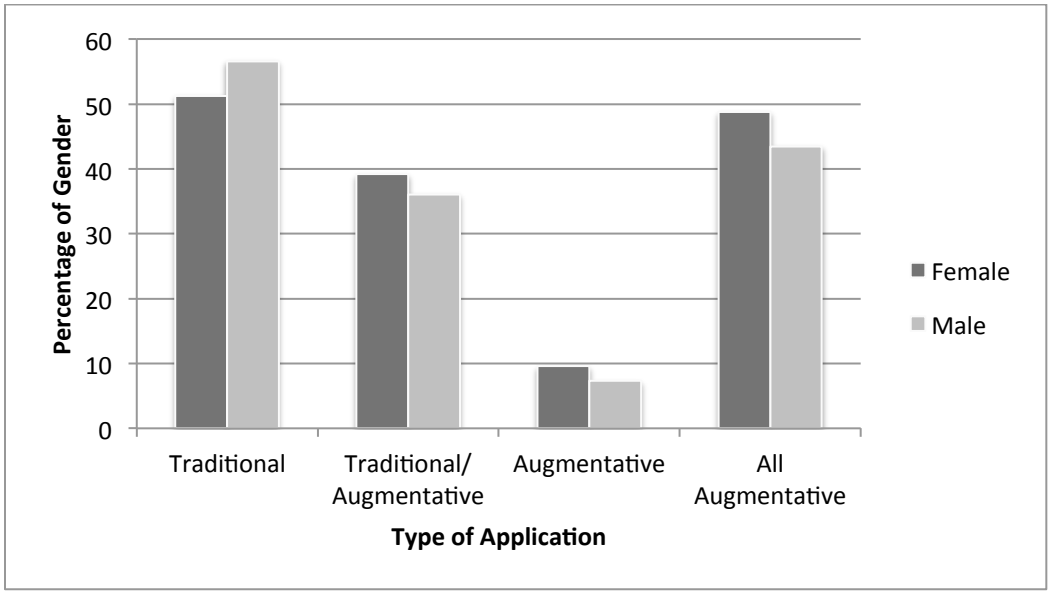
in 2012. Due to the low number of petitions for the 2012 cohort, they were removed from all statistical analysis.

The section below details the results of statistical tests for each dependent variable, with slight differences in the independent variable of application type. The first variable had three types of applications: traditional, traditional/augmentative, and augmentative. The second variable had two types: traditional, and all augmentative, which combined the latter two categories from the first variable. The third variable had two types of applications: traditional and traditional/augmentative.

Gender

The type of application submitted by males versus females in the 2012 cohort did not differ in any meaningful way (see Figure 4.9 below). It appears that a slightly higher percentage of males chose to submit traditional applications (51.25 percent of females compared to 56.56 percent of males), and a slightly lower percentage of males chose to submit augmentative applications (48.76 percent of females compared to 43.45 percent of males). The first Pearson's Chi-Square test conducted with three types of applications as the independent variable and gender

Figure 4.9 Gender Distribution for Type of Application Submitted



as the nominal dependent variable did not yield statistically significant results, X^2 (2, N=525) = 1.77, p = 0.41. The second Pearson's Chi-Square test run with only traditional and all augmentative applications also yielded a non-statistically significant result, X^2 (1, N=525) = 1.48, p = 0.22. The third Pearson's Chi-Square test run with only traditional and traditional/augmentative applications also did not yield statistically significant results, X^2 (1, N=480) = 0.94, p = 0.33. The results indicate that males and females with the option to submit different types of applications chose to submit each type of application in similar percentages; gender was not related to the type of application a student chose to submit.

Race and Ethnicity

The type of application submitted by students of different ethnicities in the 2012 cohort did not differ in any meaningful way (see Table 4.2 below). About 50 percent of students of each ethnicity submitted traditional applications, and the other 50 percent submitted augmentative applications. To avoid violating the assumption of having no cells with an expected count lower than five for the independent variable with three application types, a Pearson's Chi-Square test

Table 4.2 Ethnicity Distribution for Type of Application Submitted

	Traditional	Traditional/ Augmentative	Augmentative	All Augmentative	Total (excludes all augmentative)
African American	3	1	2	3	6
Asian American	6	3	2	5	11
Hispanic	12	9	3	12	24
Multiracial	25	14	3	17	42
Native American	8	9	1	10	18
Unknown	1	1	1	2	3
White	227	161	33	194	421
All Non-White	55	37	12	49	104

was to run with the ethnicity variable which divided all participants into non-white and white. The Pearson's Chi-Square test run with this variable did not yield statistically significant results, X^2 (2, N=525) = 1.51, p = 0.47. The second Pearson's Chi-Square test with only traditional and all augmentative application types was run with all ethnicities in the dependent variable except African American and unknown, which did not have enough responses. This test did not yield statistically significant results, X^2 (4, N=516) = 1.33, p = 0.86. The third Pearson's Chi-Square test with only traditional and traditional/augmentative applications could only be run with Hispanic, Multiracial, Native American, and White responses, to avoid violating the assumption of five expected responses per cell. This test also did not yield statistically significant results, X^2 (3, N=465) = 1.44, p = 0.70. The results indicate that students of different ethnicities with the option to submit different types of applications chose to submit each type of application in similar percentages; race or ethnicity was not related the type of application a student chose to submit.

High School Grade Point Average

The types of applications submitted by students who met the minimum HSGPA requirements to join The Honors College did not differ in a meaningful way; a student's high school grade point average did not relate to their choice to submit a traditional or traditional/augmentative application, if they were eligible to chose either based on their qualifications.

While a factorial ANOVA test with HSGPA as the dependent variable and application type and retention as the independent variables found a statistically significant result, F (2, 525) = 6.55, $p < 0.01$, this finding has no real-world significance. Students with HSGPAs above the minimum 3.75 could choose to submit either traditional (M = 3.85, SD = 0.20) or traditional/augmentative (M = 3.81, SD = 0.43) applications, while students with lower HSGPAs had no choice but to submit augmentative (M = 3.65, SD = 0.25) applications. The statistically significant difference of HSGPAs between these three application types reflects the admissions policies of The Honors College, not a free choice by the students to submit one type of application over another.

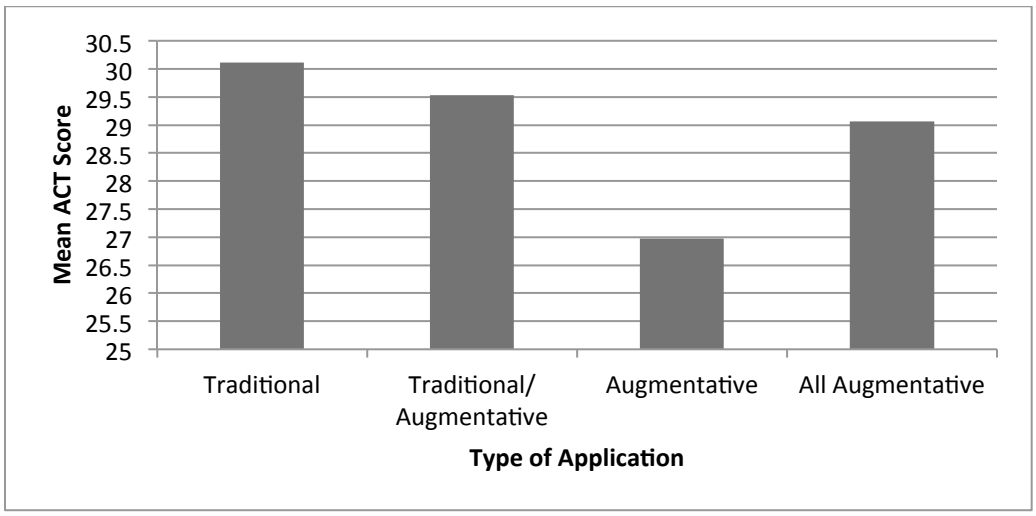
A second factorial ANOVA run comparing the HSGPAs of students who submitted traditional applications versus all augmentative (M = 3.78, SD = 0.40) applications also yielded a statistically significant result, F (1, 525) = 6.97, $p < 0.01$. However, this finding was also suspect since all augmentative applications included students who had no choice but to submit an essay in order to be admitted. This finding again reflects the policies requiring a minimum 3.75 HSGPA to submit a traditional application, rather than a free choice by students to submit any type of application.

A third factorial ANOVA run comparing the HSGPAs of students who submitted traditional versus traditional/augmentative gave the best indication of the relationship between students' high school grades and application type. These students had the free choice to submit an essay or not, due to their HSGPA meeting the minimum 3.75 requirement. This third test did not yield a statistically significant result, F (1, 480) = 2.37, p =0.12. These results indicate that students with the option to submit different types of applications based on their HSGPA chose to submit each type of application in similar percentages. High school grade point average was not related to the type of application a student chose to submit, when they were eligible to choose any type of application based on their qualifications.

ACT Scores

Among students that met the minimum 27 ACT score and 3.75 HSGPA requirements, students with lower ACT scores tended to submit traditional/augmentative applications, and students with higher ACT scores tended to submit traditional applications (see Figure 4.10 below). The types of applications submitted by students who met the minimum ACT

Figure 4.10 Mean ACT Score by Application Type



requirements did differ in a meaningful way; a student's ACT score did relate to their choice to submit a traditional or traditional/augmentative application, if they were eligible to choose either based on their qualifications.

The factorial ANOVA tests with ACT scores as the dependent variable again yielded statistically significant but meaningless results, due to the same issue discussed at length in the high school grade point average section above. The first test run with three types of applications as the independent variable yielded a statistically significant but meaningless result, F (2, 525) = 26.71, $p < 0.01$. Students with high ACT scores could choose to submit traditional (M = 30.11, SD = 2.44) or traditional/augmentative (M = 29.53, SD = 2.13) applications, while students with lower ACT scores had no choice but to submit augmentative (M = 26.98, SD = 2.67) applications.

The second test with traditional and all augmentative (M = 29.06, SD = 2.45) applications also yielded a statistically significant but meaningless result, F (1, 525) = 23.96, $p < 0.01$, since all augmentative included students with no choice but to submit an essay.

However, unlike for HSGPA, the third test comparing traditional and traditional/augmentative applications yielded a statistically significant and meaningful result, F (1, 480) = 7.15, $p < 0.01$. Levene's test for equality of error variances was met with an alpha of 0.097, indicating that variance in the two groups was equal. A Partial Eta Square value of 0.015 indicates that 1.5 percent of the variance in ACT scores is predicted by the application type the student submitted, a small effect size. These results indicate that students who could freely choose between submitting an essay or not because they met the minimum 27 ACT score and 3.75 HSGPA did choose to submit different kinds of applications. Students with higher ACT scores tended to submit traditional applications with no optional essay, while students with lower ACT scores closer to 27 tended to submit traditional/augmentative applications with an optional essay. ACT score was related to the type of application a student chose to submit, when they were eligible to choose any type of application based on their qualifications.

Parental Income Level

Students who submitted different types of applications did not have meaningfully different parental income levels. A factorial ANOVA conducted with reported parental income as the dependent variable and retention and three application types as the independent variables found that parental income was not related to type of application with statistical significance, F (2, 434) = 0.19, p = 0.83. The 231 students who submitted traditional applications (M = \$122,286.38; SD = \$103,764.01), the 166 students who submitted traditional/augmentative applications (M = \$117,697.55; SD = \$98,947.26), and the 37 students who submitted augmentative applications (M = \$115,078.43; SD = \$88,450.19) had similar parental incomes.

A second factorial ANOVA conducted with reported parental income as the dependent variable and retention and two application types (traditional and all augmentative) as the

independent variable did not yield statistically significant results, F (1, 434) = 0.122, p = 0.64. A third factorial ANOVA with two application types (traditional and traditional/augmentative) and retention as the independent variables also did not yield statistically significant results, F (1, 397) = 0.17, p = 0.68. The results indicate that students with different socioeconomic statuses, as represented by parental income, chose to submit each type of application in similar percentages. Parental income level was not related the type of application a student chose to submit.

Summary

Students submitting different kinds of applications did not differ in gender, ethnicity, high school grade point average, or reported parental income. The findings did not support the hypothesis that students who submitted augmentative applications would differ from students who submitted traditional applications, with one exception. Students in the 2012 cohort who submitted different kinds of applications to The Honors College only differed from each other in that the students who submitted traditional/augmentative applications had lower ACT scores than the students who submitted traditional applications.

The findings in this section also complicate the relationship found for part one of research question one, regarding the increase in non-white ethnicities among the 2012 cohort. Since the statistical analysis did not reveal that students of various ethnicities chose to submit different kinds of applications with any significance, the essays may not have played any role in the increase in non-white ethnicities among the 2012 cohort. The Honors College augmentative essays were based on Sternberg's WICS (Wisdom, Intelligence, Creativity, Synthesized) model (Sternberg, 2010a; Sternberg et al., 2004). Previous research regarding admissions tools based on the WICS model showed that adding optional essays led to a decrease in ethnic disparities, which may not have been observed in this research.

Effects on Retention

To test the hypothesis regarding whether students were retained differently according to which cohort they were in and which application type they submitted, Pearson's Chi-Square and

factorial ANOVA tests were run with application year or application type as the independent variable. The following section details the test results, grouped by the two dependent variables, retention and first year grade point average. Unless otherwise stated, all students in the 2011 and 2012 cohorts were included in these statistical analyses.

Retention in Second Fall and Application Type

Students' fall-to-fall retention in The Honors College was related to the type of application the students chose to submit. Students who submitted petition or augmentative applications were less likely to be retained than students who submitted traditional and traditional/augmentative applications. However, there was no meaningful difference in retention between traditional and traditional/augmentative applicants in the 2012 cohort. Figure 4.11 and Table 4.3 below illustrate the number of students admitted and retained in their second fall for the four types of applications for both cohorts. Traditional applicants were retained at 54.46 percent, compared to 32.0 percent of petition applicants, 56.57 percent of traditional/augmentative applicants, and 28.89 percent of augmentative applicants. The overall first-year retention rate for the two cohorts was 53.14 percent.

Figure 4.11 Retention by Type of Application

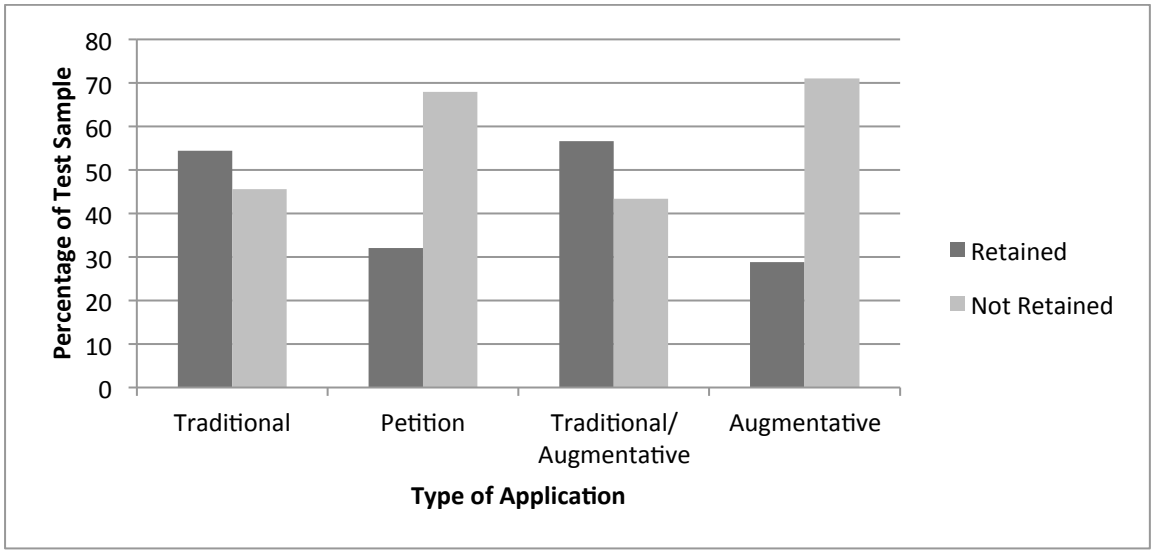


Table 4.3 Retention by Type of Application

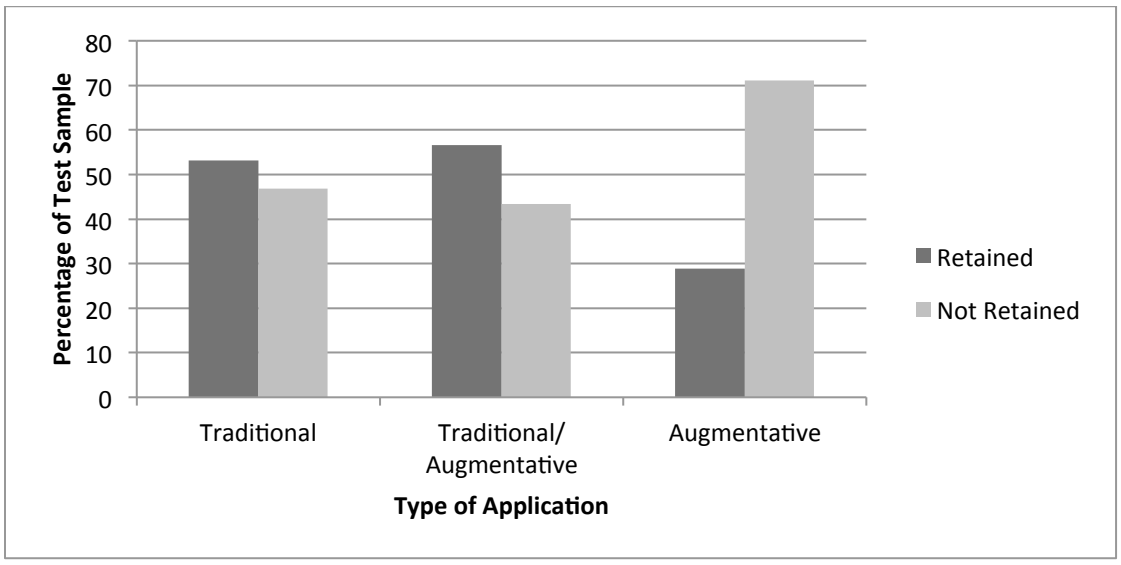
	Traditional	Petition	Traditional/ Augmentative	Augmentative	All Augmentative
Admitted	718	25	198	45	243
Retained	391	8	112	13	125
Not Retained	327	17	86	32	118

A Pearson's Chi-Square test was conducted with four types of applications (traditional, petition, traditional/augmentative, and augmentative) as the independent variable and retention as the nominal dependent variable. The results were statistically significant, X^2 (3, N=986) = 16.55, $p < 0.01$. The number of students retained who submitted petition and augmentative applications was underrepresented, with adjusted residual scores of -2.1 and -3.3, respectively. Cramer's V of 0.13 with 3 degrees of freedom indicates a small-to-medium effect size for the relationship between application type and retention. A second Pearson's Chi-Square test run with petitions removed and retention as the dependent variable also yielded statistically significant results, X^2 (2, N=961) = 11.96, $p < 0.01$, with a Cramer's V of 0.11 indicating a small effect size. The number of students retained who submitted augmentative applications was again underrepresented beyond the expected count, with an adjusted residual of -3.4. These results indicate that for both the 2011 and 2012 cohorts, whether students were retained one year later was related to the type of application the students submitted. Students who submitted petition and augmentative applications were not retained as frequently as students who submitted traditional and traditional/augmentative applications. Students who submitted petition and augmentative applications had no choice but to submit those applications because they did not meet the minimum HSGPA and ACT requirements. This finding ultimately indicates that students who do not meet the minimum HSGPA and ACT requirements are not as likely to be retained in the Honors College one year later, compared to students who do meet the requirements.

A Pearson's Chi-Square test run for only the 2012 cohort with three types of applications (traditional, traditional/augmentative, and augmentative) as the independent variables and

retention as the dependent variable also yielded statistically significant results, X^2 (2, N=525) = 11.42, $p < 0.01$, with a Cramer's V of 0.15 and 2 degrees of freedom, indicating a small-to-medium relationship for just the 2012 cohort between retention and type of application submitted. The adjusted residual for students who submitted augmentative applications and were retained was -3.3. Figure 4.12 illustrates that 53.19 percent of traditional applicants were retained,

Figure 4.12 Retention by Type of Application for 2012 Cohort



compared to 56.57 percent of traditional/augmentative, and 28.89 percent of augmentative. A second Pearson's Chi-Square conducted for only the 2012 cohort with two types of applications (traditional and traditional/augmentative) and retention as the dependent variable did not yield statistically significant results, X^2 (1, N=480) = 0.53, p = 0.47. These results indicate that for the 2012 cohort, whether students were retained one year later was related to the type of application the students submitted. Students who submitted augmentative applications were not retained as frequently as students who submitted traditional and traditional/augmentative applications. Students who submitted augmentative applications had no choice because they did not meet the minimum HSGPA and ACT requirements. This finding again indicates that students who do not

meet the minimum HSGPA and ACT requirements are not as likely to be retained in the Honors College one year later, compared to students who do meet the requirements. These findings also indicate that for students in the 2012 cohort who did meet the minimum HSGPA and ACT requirements, their choice to submit an augmentative essay or not was not related to their retention in The Honors College one year later. Students who took the time to write the optional essay were not retained any differently from students who chose not to submit an essay.

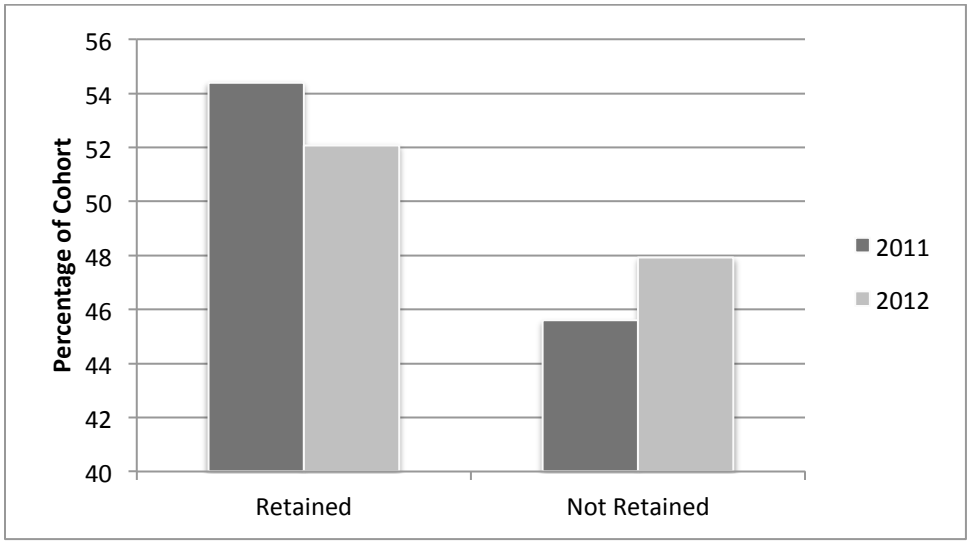
Retention in Second Fall and Cohort Year

The year in which students matriculated did not impact their retention in The Honors College. Table 4.4 below shows the number of students retained for each cohort, and Figure 4.13 illustrates that 54.39 percent of students were retained in 2011, compared to 52.08 percent in 2012. A Pearson's Chi-Square test with cohort year as the independent variable and retention as the dependent variable did not yield statistically significant results, X^2 (1, N=986) = 0.53,

Table 4.4 Retention by Cohort Year

	2011 (n=456)	2012 (n=530)	Total (N=986)
Retained	248	276	524
Not Retained	208	254	462

Figure 4.13 Retention by Cohort Year



$p = 0.47$. These findings indicate that adding the augmentative essays to the applications did not lead to a change in the fall-to-fall retention of incoming classes in The Honors College.

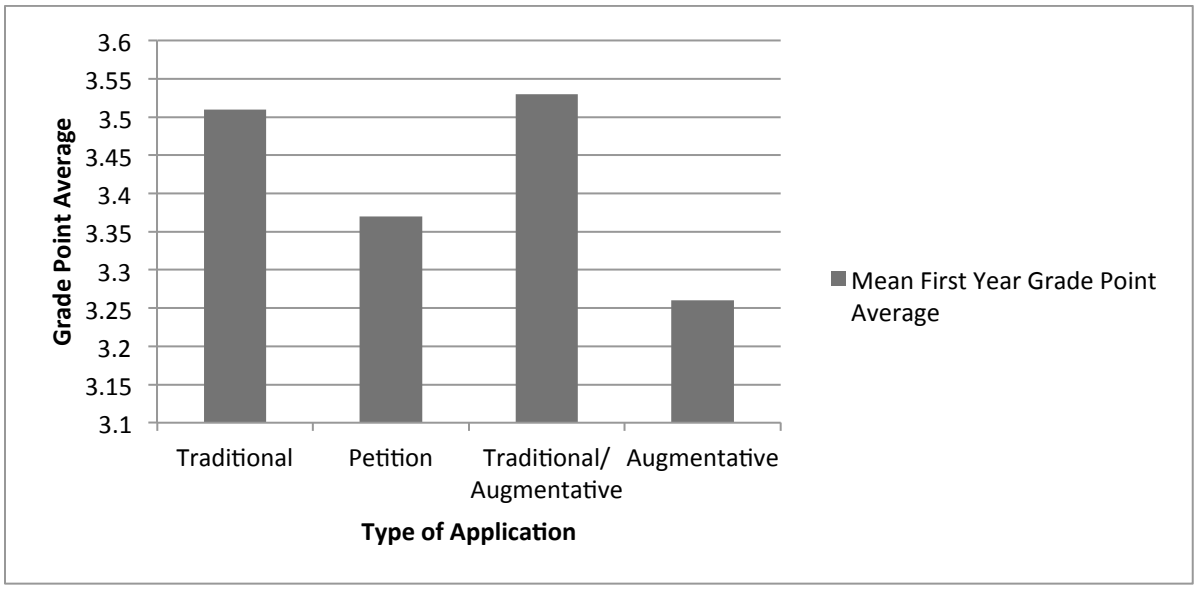
First Year Cumulative Grade Point Average and Application Type

Students admitted using augmentative applications had meaningfully lower FYGPAs than students admitted using traditional and traditional/augmentative applications. Table 4.5 and Figure 4.14 show the mean first year cumulative grade point average for each of the application

Table 4.5 Mean First Year Grade Point Averages

Cohort Year	Type of Application	Mean FYGPA	Standard Deviation	Number
2011	Traditional	3.52	0.57	436
	Petition	3.34	0.71	20
	Total	3.51	0.58	456
2012	Traditional	3.51	0.56	282
	Petition	3.49	0.44	5
	Traditional/Augmentative	3.53	0.53	198
	Augmentative	3.26	0.48	45
	Total	3.50	0.54	530
Total	Traditional	3.51	0.57	718
	Petition	3.37	0.66	25
	All Augmentative	3.48	0.53	243
	Total	3.50	0.60	986

Figure 4.14 Mean First Year Grade Point Averages by Application Type



types, for both cohorts and for all participants total. A factorial ANOVA test with was conducted with first year grade point average as the dependent variable and application type and cohort year as the independent variables. The type of application submitted proved to be statistically significant, F (3, 986) = 3.71, p = 0.01. The Partial Eta Squared of 0.011 meant 1.1 percent of the variance in FYGPA could be explained by the type of application submitted, which is a small effect size. Tukey's HSD post hoc tests revealed the two significant differences; the first was between augmentative applications and traditional applications (mean difference -0.26, p= 0.014, CI [-0.48, -0.04]), and the second was between augmentative applications and traditional/augmentative applications (mean difference -0.28, p= 0.015, CI [-0.51, -0.04]).

These findings indicate that students who applied with augmentative applications had significantly lower first year grade point averages than students who applied using traditional and traditional/augmentative applications. Again, students who applied using augmentative applications had no choice because they did not meet the minimum 3.75 HSGPA and 27 ACT requirements. The findings ultimately indicate that students who do not meet these minimum requirements earn lower first year grade point averages in their first year of college, compared to students who do meet the minimum requirements for admission to The Honors College.

First Year Cumulative Grade Point Average and Cohort Year

The various Honors College application types available in the year in which students matriculated made no meaningful difference on their first year grade point average. A one-way ANOVA conducted with FYGPA as the dependent variable and cohort year as the independent variable did not yield statistically significant results, F (1, 986) = 0.07, p = 0.79. These findings indicate that students achieved similar academic results in the 2011 and 2012 cohorts, regardless of the different application types available to each cohort.

Summary

The findings supported the hypothesis that students who submitted augmentative applications would differ from students who submitted traditional applications, and that

augmentative applicants would be retained differently than students who submitted traditional applications. Students who submitted petition and augmentative applications were less likely to be retained one year later than students who submitted traditional and traditional/augmentative applications. No difference in retention was found between students who submitted traditional and traditional/augmentative applications. In addition, students who submitted augmentative applications had lower first year grade point averages than all other students.

The students who submitted strictly augmentative applications (those who would not have been admitted without an optional essay) had lower first year grade point averages and were less likely to be retained a year later than all other application types. Augmentative applicants were retained in even lower numbers (28.89 percent) than petition applicants (32.0 percent), and had lower mean FYGPAs (3.26) than petition applicants (3.49). Augmentative applications mostly replaced petition applications among the 2012 cohort, so this surprising finding indicates that students admitted under the petition option might have additional qualities not included in this study. One possibility is that students who took the time to submit a petition, which required additional materials such as letters of recommendation, high school transcripts, and a personal statement, were more motivated to succeed than students who submitted the 400 word maximum augmentative essay. The exact differences cannot be determined by this study, but it is apparent that the augmentative applications are not an exact replacement for petition applications, and the students admitted under each option vary in important ways.

The findings also indicate that the students in the two cohorts did not meaningfully differ on retention or first year grade point average, variables which described their academic achievement after their matriculation in college, just as they did not differ on variables that described their academic achievement before matriculation in college (other than race or ethnicity). The overall 2012 cohort had similar retention rates and FYGPAs to the overall 2011 cohort, but the application types submitted within the 2012 cohort were different. Individual students, grouped by their application type, were not as successful as the overall cohort, so the

new essays may have hurt individual students most while not changing overall patterns or trends for The Honors College participants as a whole.

Factors Predicting Student Success and Retention

Unexpected findings relating to students' ACT scores and high school grade point averages necessitated further research into the predictive power of background characteristics on student success. Extant literature suggests that HSGPA may be a stronger predictor of honors student success than ACT scores, but conflicting conclusions in the field cannot fully explain the findings discussed above. Hierarchal multiple regression tests with the data for the entire sample found that high school grade point average was best able to predict both retention in honors, and first year grade point averages. While ACT scores were better able to predict first year retention than parental income level, parental income level could better predict first year grade point average than ACT scores.

A hierarchal multiple regression with retention as the dependent variable and ACT scores, HSGPA, and parental income as the independent variables found that all three independent variables were correlated with retention with statistical significance, and together accounted for 5.8 percent of the variance in whether students were retained or not ($R^2 = 0.058$, F (3, 801) = 16.32, $p < 0.01$). ACT scores alone accounted for 1.9 percent of the variance ($\beta = -0.02$, $p < 0.01$), HSGPA alone accounted for 3.3 percent of the variance ($\beta = -0.40$, $p < 0.01$), and parental income level alone accounted for 0.6 percent of the variance ($\beta < -0.01$, $p = 0.02$). These findings indicate that HSGPA was the strongest predictor of students' retention in The Honors College, but ACT scores and parental income level could also predict whether students would be retained.

A second hierarchal multiple regression with FYGPA as the dependent variable and ACT scores, HSGPA, and parental income level as the independent variables found that all three independent variables were correlated with FYGPA with statistical significance, and together accounted for 14.8 percent of the variance in first year grade point averages ($R^2 = 0.148$, F (3,

802) = 46.16, $p < 0.01$). ACT scores alone accounted for 0.9 percent of the variance ($\beta = 0.01$, $p = 0.33$), HSGPA alone accounted for 12.4 percent of the variance ($\beta = 0.87$, $p < 0.01$), and parental income alone accounted for 1.4 percent of the variance ($\beta < -0.01$, $p < 0.01$). ACT scores had a weak correlation with FYGPA of 0.097, and HSGPA had a moderate correlation with FYGPA of 0.364. These findings indicate that these three background characteristics were much stronger predictors of first year grade point average than retention in The Honors College, and that HSGPA again was the strongest predictor of first year grade point average. High school grade point average is a better indicator of students' eventual success in The Honors College than ACT scores and parental income level, though all three are significantly related to students' academic achievement and retention after their first year.

Summary of Findings

The findings chapter discussed how the data supported the hypotheses in some ways, but not in others. The race or ethnicity of students in the 2011 cohort differed from students in the 2012 cohort, though students of different ethnicities submitted each type of application in similar percentages. Students who submitted traditional/augmentative applications had lower ACT scores than students who submitted traditional applications. Students who submitted petition and augmentative applications were less likely to be retained one year later than students who submitted traditional and traditional/augmentative applications, and no difference in retention was found between students who submitted traditional and traditional/augmentative applications. In addition, students who submitted augmentative applications had lower first year grade point averages than all other students. Cohort year was not significantly related to either retention or first year grade point average. High school grade point average was the best predictor of students' first year grade point average and their fall-to-fall retention in The Honors College. The conclusion chapter will relate these findings to previous research, discuss implications for theory and practice, and suggest future directions for research.

CHAPTER V

CONCLUSION

The conclusion chapter will discuss the implications of the results of the statistical analysis. It will begin by briefly reviewing the statement of the problem, the methodology, and the results found, before discussing the results. The discussion section will include relationship to prior research, implications and recommendations for the theory, implications and recommendations for practice, and implications and recommendations for further research.

Purpose Statement

In all honors programs, diversity is an issue that should receive attention, particularly in recruitment and admission. This quantitative study examined the use of augmentative admission essays in one honors college and subsequent impacts on student diversity and retention. This research adds to the paucity of quantitative studies regarding admissions-driven efforts to increase diversity in honors programs, and contributes understanding to Sternberg's WICS model (2005) and Kaleidoscope project (2010b) in new populations.

The purpose of the present study was to examine potential differences in background characteristics and academic achievement of honors students who submitted different types of applications. Students who applied to the honors college using traditional applications (HSGPA and standardized test scores) were compared to students who applied to the honors college using augmentative applications (adding an optional essay component to the traditional application). Research questions for the study included:

1. Do students who submitted augmentative applications differ in race or ethnicity, socioeconomic status, or academic achievement from students who submitted traditional applications?

2. Are students who submitted augmentative applications retained differently than students who submitted traditional applications?

It was hypothesized that students who submitted augmentative applications would differ from students who submitted traditional applications, and that augmentative applicants would be retained differently than students who submitted traditional applications.

Review of the Methodology

An honors college at a large four-year public land-grant research institution (Carnegie Foundation for the Advancement of Teaching, 2010) in the Midwest was the location of study. The convenience-selected sample included two subgroups: the 2011-2012 entering class of The Honors College, and the 2012-2013 entering class. These subgroups only included freshmen enrolling in their first semester of college; the sample did not include transfer students or adult students with prior college credit. The sample included any student active in their first semester (enrolled in six honors hours) for the purpose of tracking retention, regardless of age, gender, ethnicity, or major. The 2011-2012 incoming class had 456 active first-semester freshmen; the 2012-2013 class had 530 active first-semester freshmen, leading to a total sample of 986 students.

The researcher collected data from two separate databases at the institution. The researcher had access to all of the necessary data, completely de-identified, and therefore did not need to obtain consent from students to undertake the present research project. Before beginning analysis, the researcher converted SAT scores to ACT scores using concordance tables available online from both College Board and ACT.

Two separate comparisons were made. The first comparison looked at two cohorts of honors students: those admitted using traditional applications in the 2011-2012 academic year, and those admitted for the following year when augmentative applications based on Sternberg's

(2005; 2010a; 2010b) research became an option for applicants. The second comparison looked only at students admitted in the 2012-2013 academic year when augmentative essays were added, but divided the students into three groups: those who submitted a traditional application without an essay, those who qualified for submitting a traditional application but chose to submit an augmentative application with an essay, and those who did not qualify for the traditional application and were admitted based on their augmentative application. The groups in both comparisons were analyzed on multiple dependent variables, including fall-to-fall retention, first-year cumulative grade point average, race or ethnicity, gender, reported parental income, standardized test scores, and high school grade point average.

The researcher analyzed the data using statistical methods in SPSS. For the interval and ratio dependent variables (FYGPA, SES, SAT/ACT scores, and HSGPA), the researcher conducted a factorial ANOVA test, with two independent variables (matriculation year, and application type). For the nominal dependent variables (retention, diversity, and gender), the researcher conducted separate Pearson's Chi-Square tests, for each of the independent variables. Statistical significance required meeting the critical p value of 0.05.

Summary of Findings

For part one of research question one, regarding the differences between students who matriculated in The Honors College in 2011 compared to 2012, only one variable, race or ethnicity, was found to have a statistically significant relationship with cohort year. Students matriculating in 2011 and 2012 did not differ in gender, high school grade point average, ACT score, or reported parental income.

For part two of research question one, regarding the differences between students in the 2012 cohort who submitted different types of applications to The Honors College, only one variable, ACT scores, was found to have a statistically significant relationship with type of application. Students submitting different kinds of applications did not differ in gender, ethnicity, high school grade point average, or reported parental income.

Research question two asked how students in the Honors College would be retained differently according to which cohort they were in and which application type they submitted. Students who submitted petition and augmentative applications were less likely to be retained one year later than students who submitted traditional and traditional/augmentative applications. No difference in retention was found between students who submitted traditional and traditional/augmentative applications. In addition, students who submitted augmentative applications had lower first year grade point averages than all other students. Findings did not indicate that cohort year was significantly related to either retention or first year grade point average.

Additional statistical analysis indicated that high school grade point average was the best predictor of both fall-to-fall retention and first year grade point average. Both ACT scores and parental income level were also statistically significant predictors of both retention and fall-to-fall retention, but with less predictive power than high school grade point average.

Relationship to Prior Research

The discussion of how the current findings relate to previous research will be divided into two sections. The first will consider how the findings relate to extant research in the field, and the second will examine how the findings relate to Sternberg's WICS (2010b) theory.

Research

The lower first year grade point averages for students who were not retained in The Honors College (M = 3.19) compared to those who were retained (M = 3.78) supports the similar findings of Cosgrove (2004) and Moon (2012), who found higher college GPAs for those students who persisted in honors compared to those who started in honors and then did not continue, and those who were invited to participate but chose not to do so. For this Honors College in particular, the lower FYGPA is not surprising because students must maintain a 3.30 institutional and cumulative grade point average after their first year to maintain eligibility to participate. Some of the students who were not retained in this honors program might have been

retained in other programs with lower continuation requirements; when staying in honors requires a higher GPA, it cannot be said with certainty that staying in honors causes a student to earn a higher GPA than they might otherwise.

The finding that high school grade point average was the strongest predictor of both fall-to-fall retention and first year grade point average adds to the paucity of quantitative empirical studies regarding the factors best able to predict success in honors programs. These findings are similar to McKay (2009), who found that high school grade point average was a better predictor for honors program completion than SAT scores, and to Bowen, Chingos, and McPherson (2009), who found that high school grade point averages alone were a stronger predictor of six-year graduation than either SAT or ACT scores. These findings contrast with Khe (2007), who could find "no consistent pattern" between high school academic achievement and college academic achievement. The present statistics also reverse the findings of Radunzel and Noble (2012), who found that ACT scores are beneficial in predicting FYGPA, but do not carry the same long-term predictive validity for student success as HSGPA. The present research indicates that ACT is less useful for predicting FYGPA than for predicting retention after one year, but this could change if the study looked instead at six year graduation rates.

In the present study, the correlation between ACT scores and FYGPA of 0.097, and the correlation between HSGPA and FYGPA of 0.364, contrasts with the findings of Kobrin, Patterson, Shaw, Mattern, and Barbut (2008), who found a correlation between SAT and FYGPA of 0.35; with Sackett, Kuncel, Arneson, Cooper, and Waters (2009), who determined that SAT scores and FYGPA have a correlation of 0.47; and with the general accepted correlation of 0.4 between SAT scores and FYGPA (Lemann, 1999). All of the studies mentioned looked at broader ranges of college students, as opposed to just high-talent honors students. It is possible that ACT scores have lower correlation when they are for the most part clustered above a score of 27, but this still does not account for the increased correlation of HSGPA with FYGPA among the high-

talent population included in the present study, especially since most HSGPAs were clustered above a 3.75 for participants.

Theory

The present study relates to previous research regarding Sternberg's (2010b) Wisdom, Intelligence, Creativity, Synthesized (WICS) model in complex ways. The lack of scores on the Honors College essays made replicating previous studies based on Sternberg's (2010b) WICS model difficult due to using categorical instead of scale variables in statistical analysis. Unlike the 2001 Rainbow Project or the 2006 Kaleidoscope Project, this research could not examine whether students' augmentative essays could increase the predictive power of the SAT or ACT on FYGPA (Sternberg, 2010a).

The present study most closely resembled the 2006 Kaleidoscope Project at Tufts University (Sternberg, 2010b). Like the present study, data for the 2011 graduates of Tufts showed the ethnic diversity of those applicants accepted to Tufts increased; black student acceptance increased 30% in the first year of Kaleidoscope, and Hispanic student acceptance increased 15% in the first year of Kaleidoscope. The Kaleidoscope Project included all applicants, whether they chose to matriculate at Tufts or not. The present study did not look at students who were accepted but chose not to participate in the Honors College, but the ethnic diversity of the 2012 cohort did differ significantly from the ethnic diversity of the 2011 cohort. These findings are also similar to the University of Michigan Business Project, which found that the newly developed questions reduced gender and ethnic disparities (Sternberg, The Rainbow Project Collaborators, and the University of Michigan Business School Project Collaborators, 2004), and the Rainbow Project, which found that the new tests reduced the deviation from the mean group differences among students from non-white ethnic backgrounds (Sternberg, 2010a).

The statistical analysis of the 2012 cohort alone did not find that students who wrote the optional essays differed in any way from the students who did not write the optional essays, including FYGPA, which contrasts with the findings from Kaleidoscope, wherein students who

wrote the optional essays were more likely to have higher FYGPAs (Sternberg, 2010b). The University of Michigan Business School Project also found that students who performed well on questions developed based on the WICS model had higher FYGPAs (Sternberg et al., 2004). The obvious difference is that all 422 students at the University of Michigan took the tests involving analytical, practical, and creative intelligence after they had been admitted, not as a condition of admission.

Implications and Recommendations for Further Research

There are many directions for future research with this project. A continuing comparison of the students admitted under each application type through their eventual graduation would be even more insightful for understanding how the cohorts are retained throughout their entire undergraduate career (Bowen, Chingos, & McPherson, 2009). Many students in The Honors College stop participating after they finish the first of three separate awards, especially if they are concerned about the difficulty of a senior honors thesis. It would be interesting and beneficial to examine how many of the traditional, traditional/augmentative, and augmentative applicants go on to finish the thesis and earn the Honors College Degree, compared to previous years in which students only had the option of applying with traditional or petition applications.

Expanding the study to include multiple years of students before the essay applications and multiple years after would give an ever stronger indication of changes within the Honors College population over time. There are several nuances within the 2011 and 2012 cohorts which alone could have led to significant findings which were Type I errors, or which could have led to non-statistically significant findings which were Type II errors (Nolan & Heinzen, 2012). Replicating the study with more participants could lead to more definitive patterns and findings.

Additional variables could be studied to discover other differences between augmentative and traditional applicants, and between traditional/augmentative and traditional applicants. First generation status of students and the highest level of parental education were included in

statistical analysis for the current study, but no statistically significant results were found for either. Additional variables beyond those could also include AP scores and high school rank.

Another direction for future study within this same honors college would be to compare those students admitted in the spring semester of their freshman year to augmentative and traditional students. Spring admits sometimes were unaware of the Honors College before they started their undergraduate careers, or they were concerned about its difficulty and their ability to successfully participate. An invitation sent out to all first-semester freshmen earning a 3.50 or above in their first semester generally yields great interest in students who would not have met the traditional application requirements coming straight out of high school, and oftentimes those students are retained in high numbers.

Qualitative studies regarding the motivation of students to submit certain kinds of essay applications could also help The Honors College understand possible differences in applicants that are not discernible in quantitative analysis. Sternberg's (2010b) research is based on the assumption that individuals have different kinds of intelligence, but the only way to examine the kinds of intelligence among Honors College participants would be to work more closely with students using interviews and the kinds of tests employed in the Rainbow Project (2010a).

Finally, the lower ACT scores among students submitting traditional/augmentative applications, compared to students submitting traditional applications, was an unexpected finding. An ACT score of 27 of higher qualified students to submit a traditional application, and students could read that they qualified for this option on the application website. Students with a score close to or exactly at 27 may have believed that submitting an essay along with their application would help their chances of being admitted, if they did not know that they would be automatically accepted either way. Students with very high ACT scores may have felt more confident in their applications without submitting an optional essay. Alternative and more complicated explanations are also possible, such as a high score on the ACT correlating with lower motivation in college applications.

Future research is needed to continue exploring the complex relationships among these variables, especially for honors and high-talent students.

Implications and Recommendations for the Theory

The present findings may reinforce the ability of Sternberg's (2010b) WICS model to reduce ethnic disparities among students in various educational settings. Sternberg believes students from different ethnic or cultural groups have different kinds of intelligences based on their cultural values, experiences, norms, and challenges. Essays, tests, or other admissions tools that allow students to showcase unique intelligences (whether their strengths lie in wisdom, analytical, practical, or creative intelligences) provide admissions representatives a better way to compare and admit students than standardized tests that only show analytical intelligence. The Honors College augmentative essays may not have directly caused the increase in non-white students among the 2012 cohort, but the increase was significant in the first year that essays were an available option.

The results of the present study may suggest that the WICS model needs slight revision. The augmentative essays in this study were intended to discover wisdom, analytical intelligence, practical intelligence, and creativity among the incoming students, traits which Sternberg maintains are not evident from students' high school grade point averages and especially their standardized test scores. Students admitted to The Honors College in the augmentative category did not meet the minimum HSGPA and ACT requirements but did write strong essays showcasing other talents; these students therefore represent the students Sternberg's theory aims to discover and nourish. Belonging to The Honors College brings students additional benefits, such as increased contact with faculty, smaller class sizes, additional academic resources, and more discussion-based classrooms that value creativity. It seems reasonable that the augmentative applicants would thrive in this environment, yet they were retained in the lowest percentage, compared to students admitted based on their higher HSGPA and ACT scores. It is possible, as Sternberg (2010b) writes, that this disconnect arises from educational classrooms continuing to

value the rote memorization which allows certain students to excel on standardized tests in the first place. Yet the purpose of The Honors College is to provide highly talented students with the opposite academic milieu, the kind Sternberg (2010b) advocates creating. The findings therefore may suggest that the WICS model may be successful in initially discovering students with different kinds of intelligence, but that those students do not excel in high talent honors classrooms on the basis of those WICS talents alone. If the students identified as future leaders and creative thinkers by the WICS model do not succeed in the best possible academic environment available to them, the WICS theory needs to further address why individuals who are successful later in life cannot thrive even in a specialized academic setting.

The results also suggest that the admissions methods based on the WICS model used to find students of different intelligences could be reconsidered. The findings from the Rainbow Project, the most developed of the tests based on Sternberg's theory, were most ground-breaking in reducing ethnic and gender disparities, and enhancing the predictive ability of the SAT. The findings from Project Kaleidoscope, the optional essays implemented at Tufts University, were less remarkable in reducing ethnic disparities and enhancing the predictive ability of the SAT. Despite the statistical challenges in the present study, findings indicate the optional essays used by The Honors College were even less successful in producing the expected changes among the student body. The further the admissions tools based on the WICS model have evolved from the original Rainbow Project, the less beneficial they seem to be. This does not mean the WICS model is not useful; it means the methods used to test and implement the WICS model need to be reevaluated. Optional essays as a condition for admission among high-talent students do not appear to yield promising results for discovering and considering people's inherent diverse strengths.

Implications and Recommendations for Practice

Based on the findings, The Honors College at this institution may want to consider other admissions methods to recruit and retain students. Using the optional essays to admit students

with HSGPAs and ACT scores lower than the traditional requirement may not be in the best interest of the students, and may be a waste of time and resources on the part of The Honors College. Approximately one in four augmentative applicants were retained one year later. Augmentative students may have been better served by waiting to join The Honors College in the spring semester of their freshman year, based on their first semester grade point averages, when their ability to achieve academic success in college was more solidified. The Honors College could also reconsider how the optional essays are used. Sternberg (2010b) only started using the WICS-derived essays as an admissions tool at Tufts University in 2006, and a scoring system was devised. Perhaps The Honors College could discontinue the minimum ACT or SAT scores required for traditional acceptance, and could encourage more students to submit the optional essays, which would be scored instead of voted over. The ACT scores account for very little of students' eventual retention or FYGPA, and if the essays were scored as opposed to voted over, Honors College staff may find results more consistent with Sternberg's earlier work.

Other honors programs in the United States should still consider Sternberg's WICS model (2010b) as one possible answer to the difficulty of admitting diverse qualified high-talent students. The increase in non-white ethnicities among the 2012 cohort is promising, especially given the lack of representation in some honors programs for non-white students. Other honors programs should also note the usefulness of HSGPA in predicting first year retention and FYGPA, compared to ACT scores, in considering what criteria should be used to admit high-talent students.

Conclusion

This research study was intended to increase understanding of how to recruit and retain diverse students in honors programs, particularly through the use of alternative admissions policies. Honors administrators have a responsibility to admit qualified students from all backgrounds, so that they may benefit from the many increased learning outcomes of honors participation. This study provides insight into one potential method, augmentative essays

designed to elicit unique intelligences, administrators may use to recruit diverse students. The first year in which students had the option to submit augmentative applications to The Honors College, there was a significant increase in non-white students. While augmentative essays may not be the best potential method for admitting diverse honors students, this study provides a starting place other programs can use in exploring admissions tools. Each honors program is unique, but the best way for these programs to progress and serve all students is to learn from the mistakes and successes of other honors programs.

In addition, this study joins the debate over the use of standardized test scores in university admissions. The findings add to the limited quantitative understanding of which factors best predict success in honors programs; high school grade point average was shown to be a better predictor of retention and first year grade point average than ACT scores. Numerous studies indicate that the reliance of universities on standardized test scores may harm students from non-white and low socioeconomic backgrounds. The combined insights that the augmentative essays may have led to an increase in non-white students, and that high school grade point average predicts success in honors better than ACT scores, suggest that honors programs should move away from using minimum ACT scores as part of their admissions criteria, especially to increase the acceptance rate of diverse students.

Current national trends towards quantitative assessment of higher education raise the alarming possibility that administrators will rely more on ACT scores or other similar standardized tests. Doing so would continue to harm both talented students who could benefit from participating in honors programs, and society, which could benefit from the talents of honors-educated students in the future. Every effort to explore holistic application procedures for honors programs, especially for large programs without the resources to utilize individual interviews, brings the educational community one step closer to finding a solution to recruit, admit, and retain qualified and diverse honors students.

REFERENCES

ACT (2008). ACT – SAT concordance. Retrieved March 24, 2013, from http://www.act.org/aap/concordance/pdf/reference.pdf

Andrews, A. (2011). The wisdom of our elders: Honors discussions in *The Superior Student*, 1958-1965. *Journal of the National Collegiate Honors Council, 12*(2): 17-46.

Atkinson, R. C., & Geiser, S. (2009). Reflections on a century of college admissions tests. *Center for Studies in Higher Education*, *4*(9): 1-21. Retrieved September 23, 2012 from http://escholarship.org/uc/item/49z7127p

Austin, C. G. (1986). Orientation to honors education. In P. G. Friedman & R. C. Jenkins-Freidman (Eds.), *Fostering academic excellence through honors programs*, 5-16. San Francisco: Jossey-Bass Publishers.

Aydelotte, F. (1925). Bulletin of the National Research Council: Honors courses in American colleges and universities (2nd ed.). Washington, D.C.: The National Research Council of the National Academy of Sciences.

Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line: Completing college at America's public universities*. Princeton: Princeton University Press.

Carnegie Foundation for the Advancement of Teaching (2010). Carnegie classification of institutions of higher education. Retrieved March 24, 2013, from http://classifications.carnegiefoundation.org/

College Board (2009). ACT and SAT concordance tables. Retrieved March 24, 2013, from

http://research.collegeboard.org/publications/content/2012/05/act-and-sat-concordance-tables

College Board (2012). About the SAT. Retrieved July 9, 2012, from http://sat.collegeboard.org/why-sat/topic/sat/what-the-sat-tests

Cosgrove, J. (2004). The impact of honors programs on undergraduate academic performance, retention, and graduation. *The Journal of the National Collegiate Honors Council 5*(2), 45-53. Retrieved from http://digitalcommons.unl.edu/nchcjournal/137/

Espenshade, T. J., Hale, L. E., & Chung, C. Y. (2005). The frog pond revisited: High school academic context, class rank, and elite college admission. *Sociology of Education, 78*(4), 269-293. Retrieved from http://www.jstor.org/stable/4150499

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the Scholastic Assessment Test and general cognitive ability. *Psychological Science, 15*, 373-378. doi:10.1111/j.0956-7976.2004.00687.x

Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences* (4th ed.). New York, NY: Basic Books.

Gay, L. R., Mills, G. E., & Airasian, P. W. (2011). *Educational research: Competencies for analysis and applications* (10th ed.). Upper Saddle River, NJ: Pearson Education.

Geiser, S., & Studley, R. (2002). UC and the SAT: Predictive validity and differential impact of the SAT I and SAT II at the University of California. *Educational Assessment, 8*(1): 1-26. doi: 10.1207/S15326977EA0801_01

Gratz v. Bollinger, 539 U.S. 244 (2003).

Gurin, P., Dey, E. L., Hurtado, S., & Gurin, G. (2002). Diversity and higher education: Theory and impact on educational outcomes. *Harvard Educational Review, 72*(3): 330-366.

Honors College optional freshman admission essay questions: 2012-2013. (2011). Retrieved June 25, 2012.

Hopwood v. Texas, 78 F.3d 932 (5th Cir. 1996).

Khe, S. (2007). The irrelevance of SAT in honors? *Journal of the National Collegiate Honors Council, 8*(2), 79-86. Retrieved from http://digitalcommons.unl.edu/nchcjournal/48/

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. The College Board Research Report No. 2008-5. Retrieved from http://research.collegeboard.org/publications/content/2012/05/validity-sat-predicting-first-year-college-grade-point-average

Lemann, N. (1999). *The big test: The secret history of the American meritocracy*. New York: Farrar, Straus and Giroux.

Long, B. T. (2002). Attracting the best: The use of honors programs to compete for students. Harvard University: unpublished working paper for the Spencer Foundation. Retrieved from http://gseacademic.harvard.edu/~longbr/HonorsColleges.pdf

McKay, K. (2009). Predicting retention in honors programs. *Journal of the National Collegiate Honors Council, 10*, 77-88. Retrieved from http://digitalcommons.unl.edu/nchcjournal/253/

Mohler, M. (2012a). The growth of honors programs in the 20th century: One university's experience. Unpublished manuscript.

Mohler, M. (2012b). The rise of standardized tests: Admissions changes in American colleges during the 20th century. Unpublished manuscript.

Moon, J. L. (2012). Honors and high-ability students: Factors that predict academic efficacy, critical thinking skills, and academic goals. *Dissertation Abstracts International,* 3511628.

National Association for College Admission Counseling (2008). Report of the commission on the use of standardized tests in undergraduate admission. Retrieved from http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED502721

National Collegiate Honors Council (2010a). Basic characteristics of a fully developed honors college. Retrieved March 24, 2013, from http://nchchonors.org/faculty-directors/basic-characteristics-of-a-fully-developed-honors-college/

National Collegiate Honors Council (2010b). Basic characteristics of a fully developed honors program. Retrieved March 24, 2013, from http://nchchonors.org/faculty-directors/basic-characteristics-of-a-fully-developed-honors-program/

Nolan, S. A., & Heinzen, T. E. (2012). *Statistics for the behavioral sciences* (2nd ed.). New York, NY: Worth Publishers.

Northouse, P. G. (2010). *Leadership: Theory and practice* (5th ed.). Thousand Oaks, CA: Sage Publications, Inc.

Present Student Body (2012). Retrieved March 25, 2013 from http://vpaf.okstate.edu/irim/StudentProfile/2012/PDF/2012PSB.pdf

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. ACT Research Report Series 2012(5). Retrieved from http://media.act.org/documents/ACT_RR2012-5.pdf

Regents of the University of California v. Bakke, 438 U.S. 265 (1978).

Rinn, A. (2006). Major forerunners to honors education at the collegiate level. *Journal of the National Collegiate Honors Council, 7*(2), 63-84. Retrieved from http://digitalcommons.unl.edu/nchcjournal/17/

Rinn, A. N., & Cobane, C. T. (2009). Elitism misunderstood: In defense of equal opportunity. *The Journal of the National Collegiate Honors council 10*(1), 53.56. Retrieved from http://digitalcommons.unl.edu/nchcjournal/256/

Rothstein, J. M. (2004). College performance predictions and the SAT. *Journal of Econometrics, 121*, 297-317. doi: 10.1016/j.jeconom.2003.10.003

Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009).

Socioeconomic status and the relationship between the SAT and freshman GPA: An analysis of data from 41 colleges and universities. The College Board Research Report 2009-1. Retrieved from http://research.collegeboard.org/sites/default/files/publications/2012/9/researchreport-2009-1-socioeconomic-status-sat-freshman-gpa-analysis-data.pdf

Schuman, S. (2006). Beginning in honors: A handbook (4th ed.). *NCHC Monograph Series, 7*, 1-80. Retrieved from http://digitalcommons.unl.edu/nchcmono/7/

Seifert, T. A., Pascarella, E. T., Colangelo, N., & Assouline, S. (2007). The effects of honors program participation on experiences of good practices and learning outcomes. *Journal of College Student Development, 48*(1): 57-74. doi: 10.1353/csd.2007.0007

Selingo, J. (2002, May 31). Honors programs boom at community colleges. *The Chronicle of Higher Education*. Retrieved July 18, 2012 from http://chronicle.com/article/Honors-Programs-Boom-at/33564/

Shushok, F. (2002). Educating the best and the brightest: Collegiate honors programs and the intellectual, social, and psychological development of students. *Dissertation Abstracts International,* 3070562.

Sternberg, R. J. (2005). WICS: A model of leadership. *The Psychologist-Manager Journal, 8*(1), 29-43. doi: 10.1207/s15503461tpmj0801_4

Sternberg, R. J. (2010a). Assessment of gifted students for identification purposes: New techniques for a new millennium. *Learning and Individual Differences, 20*(4), 327-336. doi: 10.1016/j.lindif.2009.08.003

Sternberg, R. J. (2010b). *College admissions for the 21st century*. Cambridge: Harvard University Press.

Sternberg, R. J., Ferrari, M., Clinkenbeard, P. R., & Grigorenko, E. L. (1996). Identification, instruction, and assessment of gifted children: A construct validation of a triarchic model. *Gifted Child Quarterly, 40*, 129-137. doi: 10.1177/001698629604000303

Sternberg, R. J., & The Rainbow Project Collaborators (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence, 34*, 321-350. Retrieved July 18, 2012 from http://www3.wooster.edu/teagle/docs/The%20Rainbow%20Project.pdf

Sternberg, R. J., The Rainbow Project Collaborators, & The University of Michigan Business School Project Collaborators (2004). Theory-based university admissions testing for a new millennium. *Educational Psychologist, 39*(3): 185- 198. doi: 10.1207/s15326985ep3903_4

Thelin, J. R., & Hirschy, A. S. (2009). College students and the curriculum: The fantastic voyage of higher education, 1636 to present. *The Journal of the National Academic Advising Association, 29*(2): 9-17.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: University of Chicago Press.

Zwick, R. (2007). College admissions in twenty-first-century America: The role of grades, tests, and games of chance. *Harvard Educational Review 77*(4), 419-429. Retrieved from http://her.hepg.org/content/u67n84589527t80v/

Zwick, R., & Green, J. G. (2007). New perspectives on the correlation of SAT scores, high school grades, and socioeconomic factors. *Journal of Educational Measurement, 44*, 23-45. doi: 10.1111/j.1745-3984.2007.00025.x

APPENDICES

Appendix 1

Optional Essay Questions for the 2012-2013 incoming freshmen:

1. One of your best friends tells you in strict confidence that he is cheating on his girlfriend who also is one of your best friends. Now the girlfriend has come to you because she is suspicious that her boyfriend is cheating on her. What should you do? Why? [Feel free to change the question to a girlfriend who is cheating on her boyfriend.]

2. You have just learned that the world will lose all electronic communication for a ten-year period beginning August 15, 2012, because of massive cosmic energy storms. How will society be affected (for ill or for good, or both), and what will you do under these circumstances?

3. According to Winston Churchill, "Criticism may not be agreeable, but it is necessary. It fulfills the same function as pain in the human body. It calls attention to an unhealthy state of things." Do you find Churchill's observation to be true in the 21st Century? Explain why or why not.

4. Write a short story using one of the following titles: Betrayed! or Oops! or You're Kidding??!! or The Last Day.

5. Sometimes "standard" questions like those above do not allow you to express yourself adequately. If you believe this to be the case, develop your own question and answer it thoughtfully (“Honors College Optional Freshman Admission Essay Questions,” 2011).

Oklahoma State University Institutional Review Board

Request for Determination of Non-Research or Non-Human Subject

Federal regulations and OSU policy require IRB review of all research involving human subjects. Some categories of research are difficult to discern as to whether they qualify as human subject research. Therefore, the IRB has established policies and procedures to assist in this determination.

1. Principal Investigator Information

First Name: McKenzie	Middle Initial: L.	Last Name: Mohler
Department/Division: Educational Leadership Studies		College: OSU College of Education
Campus Address: 101 Old Central		Zip+4: 74078-1073
Campus Phone: 405-744-6768	Fax: 405-744-6839	Email: mckenzie.mohler@okstate.edu
Complete if PI does not have campus address:		
Address:		City:
State:	Zip:	Phone:

2. Faculty Advisor (complete if PI is a student, resident, or fellow) ☐ **NA**

Faculty Advisor's name: Tami Moore		Title: Assistant Professor
Department/Division: Educational Studies		College: College of Education
Campus Address: 700 Greenwood Ave. / Main Hall, Room 2439 / Tulsa, OK 74106		Zip+4:
Campus Phone: 918-594-8107	Fax:	Email: tami.moore@okstate.edu

3. Study Information:

A. Title

Supplementing Honors College Applications with an Optional Essay: Effects on Cohort Differences, First Year Retention, and Ethnic and Gender Disparities

B. Give a brief summary of the project. (See instructions for guidance)

The proposed research project is a quantitative study to examine the use of augmentative admission essays in The Honors College and subsequent impacts on study diversity and retention. The purpose of the study is to examine potential differences between honors students based on type of application. Students who applied to the honors college using traditional applications (HSGPA and standardized test scores) will be compared to students who applied to the honors college using augmentative applications (adding an optional essay component to the traditional applications). Research questions for the study include: 1) Do students who submitted augmentative applications differ in race or ethnicity, socioeconomic status, or academic achievement from students who submitted traditional applications?; and 2) Are students who submitted augmentative applications retained differently than students who submitted traditional applications?

This study does not fit into one particular subgroup of quantitative research; it blends causal-comparative research and correlational research, in the sense that it seeks to explain relationships or patterns using secondary analysis of pre-existing data, but without determining causation. The design will allow the

research to determine what, if any, differences exist between two matriculating classes in The Honors College. Any statistically significant relationships discovered will require further research using experimental designs if possible, to determine causation.

The pre-existing data will be compiled from two separate databases on campus (described in further detail below). The researcher will receive de-identified data for statistical analysis, including factorial ANOVA and Chi-Square tests.

C. Describe the subject population/type of data/specimens to be studied. (See instructions for guidance)

The population for this research encompasses all honors students at the institution studied. The Honors College typically includes 5-6 percent of the entire undergraduate body. This population does not include those who apply to participate, but either do not gain admission or choose not to participate in The Honors College. The sample for this study will be chosen using convenience sampling; the sample will include all of the individuals within two consecutive incoming classes of The Honors College. The sample will only include freshmen (typically 18+ years old) enrolling in their first semester of college; the sample will not include transfer students or adults with prior college credit. The total sample size is 990 students (458 from the 2011-2012 incoming class, and 532 from the 2012-2013 incoming class).

The researcher will collect data from two separate databases at the institution. The institution collects data for SAT and ACT scores, race or ethnicity, gender, reported parental income, FYGPA, and HSGPA in the campus-wide database Student Information Systems (SIS). The Honors College keeps a separate Filemaker database which houses information on whether students submitted optional essays, whether students met normal admission requirements, and active status for each semester. To compile data from these two sources, first an Honors College employee other than the researcher will export the necessary data from Filemaker into an Excel spreadsheet, along with CWIDs (Campus Wide Identification numbers), and assign randomized numbers found online using a website such as random.org. The employee will then send this spreadsheet to the Institutional Research and Information Management (IRIM) office on campus. IRIM will match the information from Filemaker with the information from SIS, de-identify the data by deleting all CWIDs, and will send the spreadsheet back to the researcher. The researcher will then have access to all of the necessary data, completely de-identified, and will therefore not need to obtain consent from students to undertake the present research project. Before beginning analysis, the researcher will convert SAT and ACT scores to the same scale using concordance tables available from both College Board and ACT online. Since ACT is more prevalent for in-state students, SAT scores will be converted to the ACT scale.
The researcher will begin analyzing data for the 2011 entering class in summer 2013, and will begin analyzing the data for the 2012 entering class after the third week of the fall 2013 semester, at which point The Honors College will have determined active status, and thus fall-to-fall retention, for the entering class of fall 2012. This will necessitate two separate spreadsheets sent to IRIM for data compilation, with distinct randomized numbers to eliminate confusion.

4. Determination of "Research".
45 CFR 46.102(d): *Research* means a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge. Activities which meet this definition constitute research for purposes of this policy whether or not they are conducted or supported under a program which is considered research for other purposes.

One of the following must be "no" to qualify as "non-research":

A. Will the data/specimen(s) be obtained in a systematic manner?
☐ No ☒ Yes

B. Will the intent of the data/specimen collection be for the purpose of contributing to generalizable knowledge (the results (or conclusions) of the activity are intended to be extended beyond a single individual or an internal program, e.g., publications or presentations)?
☐ No ☒ Yes

Oklahoma State University Institutional Review Board

Request for Determination of Non-Research or Non-Human Subject

5. Determination of "Human Subject".

45 CFR 46.102(f): *Human subject* means a living individual about whom an investigator (whether professional or student) conducting research obtains: (1) data through intervention or interaction with the individual or (2) identifiable private information. Intervention includes both physical procedures by which data are gathered (for example venipuncture) and manipulations of the subject or the subject's environment that are performed for research purposes. Interaction includes communication or interpersonal contact between investigator and subject. Private information includes information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, and information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (for example, a medical record). Private information must be individually identifiable (i.e., the identity of the subject is or may be ascertained by the investigator or associated with the information) in order for obtaining the information to constitute research involving human subjects.

A. Does the research involve obtaining information about living individuals?
☐ No ☒ Yes
If no, then research does not involve human subjects, no other information is required.
If yes, proceed to the following questions.

All of the following must be "no" to qualify as "non-human subject":

B. Does the study involve intervention or interaction with a "human subject"?
☒ No ☐ Yes

C. Does the study involve access to identifiable private information?
☒ No ☐ Yes

D. Are data/specimens received by the Investigator with identifiable private information?
☒ No ☐ Yes

E. Are the data/specimen(s) coded such that a link exists that could allow the data/specimen(s) to be re-identified?
☒ No ☐ Yes
If "*Yes*," is there a written agreement that prohibits the PI and his/her staff access to the link?
☐ No ☐ Yes

6. Signatures

Signature of PI _____ **Date** 4/24/13

Signature of Faculty Advisor (If PI is a student) _____ **Date** 4/24/13

☒ Based on the information provided, the OSU-Stillwater IRB has determined that this project **does not** qualify as human subject research as defined in 45 CFR 46.102(d) and (f) and **is not subject to oversight by the OSU IRB.**

☐ Based on the information provided, the OSU-Stillwater IRB has determined that this research **does** qualify as human subject research and **submission of an application for review by the IRB is required.**

Shelia M Kennison
Dr. Shelia Kennison, IRB Chair

5-6-13
Date

VITA

McKenzie Leigh Mohler

Candidate for the Degree of

Master of Science

Thesis: SUPPLEMENTING HONORS COLLEGE APPLICATIONS WITH AN OPTIONAL ESSAY: EFFECTS ON COHORT DIFFERENCES, FIRST YEAR RETENTION, AND ETHNIC AND GENDER DISPARITIES

Major Field: Educational Leadership Studies, Option in Higher Education Administration

Biographical:

Education:

Completed the requirements for the Master of Science in Educational Leadership Studies at Oklahoma State University, Stillwater, Oklahoma in December 2013.

Completed the requirements for the Bachelor of Arts in English Literary Studies at Washington State University, Pullman, Washington in May 2010.

Experience:

Honors Academic Counselor at The Honors College of Oklahoma State University from August 2010 to present.

Intern at the Oklahoma State Regents of Higher Education from January 2013 to May 2013.

Orientation Counselor at New Student Programs of Washington State University from November 2008 to August 2010.

Administrative Tutor / Tutor Legacy Fellow at the Writing Center of Washington State University from August 2008 to May 2010.

Professional Memberships:

National Collegiate Honors Council