MISSOURI GENERAL EDUCATION ASSESSMENT:
EXAMINATION OF TEACHER CANDIDATE SCORES AND PREDICTOR VARIABLES

A Dissertation presented to
the Faculty of the Graduate School
University of Missouri – Columbia

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Education

by
JEFFREY M. EDMONDS
Dr. Timothy J. Wall, Dissertation Supervisor

DECEMBER 2014
The undersigned, appointed by the dean of the Graduate School,

have examined the dissertation entitled

MISSOURI GENERAL EDUCATION ASSESSMENT:
EXAMINATION OF TEACHER CANDIDATE SCORES AND PREDICTOR VARIABLES

presented by Jeffrey M. Edmonds,
a candidate for the degree of

Doctor of Education

and hereby certify that, in their opinion,

it is worthy of acceptance.

______________________________
Dr. Timothy J. Wall, Dissertation Supervisor

______________________________
Dr. Matt Symonds, Committee Member

______________________________
Dr. Joyce Piveral, Committee Member

______________________________
Dr. Jeffrey Foot, Committee Member
ACKNOWLEDGEMENTS

I want to first thank my mother, Dr. Carole Edmonds. She is my hero and primary role model and was a member of the original cohort in the statewide ELPA program. Mom, I want to be you when I grow up. I can’t thank you enough for being my unofficial advisor throughout this process. It is safe to say I could not have done this without you. It should be noted that I probably would not have done this without you.

I also want to thank a number of my friends. Rob Duvall and Sarah Thomas, thanks for being editors of my final drafts and for getting me over the last hurdle of this process. Sarah, you and your life at the time of my final push humbled me beyond words. This is the only section of my dissertation in which my thoughts can’t be summarized. Chad Harris, thanks for being my Cohort 8 partner-in-crime. I could not have asked for a more like-minded academic classmate, and your friendship and camaraderie made our two years of coursework both fruitful and enjoyable. Megan van Alstine, thanks for being my unofficial partner-in-crime. Once you finally finish your “doctorate”, we can travel the world again. Rob Smitherman and James Tesdall, thanks for being awesome roommates during my coursework and allowing me to “attend” class via satellite while you were living your lives in the dining room, living room, or wherever you might have been on Wednesday nights. Jeannie Kenyon, Adam Cartwright, and Mitch and Megan Ronnei, thanks for always reminding me I was doing something important and substantial and that I should be proud of it.

Additional thanks are owed to Ms. Cory Stutts and the Catherine Cook School of Chicago, Illinois. Cory, thank you for allowing me the time to pursue this degree, especially when it wasn’t convenient. Thanks to my colleagues—Sara, Jessica, Kate,
Toni, and Lara—for covering for me at important times. You made this whole thing possible by making small but critical sacrifices.

Finally, I want to thank Northwest Missouri State University personnel. Mike McBride, thanks for being my third-party data collector. I honestly could not have done this without you. Seriously, it wouldn’t have been possible. Dr. Piveral and Dr. Foote, thank you for serving on my committee. Your feedback at the time of my proposal was quite helpful and appreciated. Dr. Wall, thanks for being my advisor throughout this process. I greatly appreciated your guidance and support. Dr. Symonds, thanks for the emails, text messages, and phone calls. Your support was critical to this process, and I can’t thank you enough for giving up your free time to help me on numerous occasions. Dr. Wall and Dr. Symonds, I feel we did this very quickly, but I’m confident we also did it very skillfully and thoughtfully.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................ iv

LIST OF FIGURES .................................................................................................................. x

LIST OF TABLES ..................................................................................................................... xii

ABSTRACT .............................................................................................................................. xiv

CHAPTER ONE: INTRODUCTION TO THE STUDY .............................................................. 1

  Background ....................................................................................................................... 2

  Conceptual Underpinnings ............................................................................................... 9

  Problem Statement ....................................................................................................... 12

  Research Purpose ....................................................................................................... 13

  Research Questions and Null Hypotheses .................................................................. 13

  Limitations, Delimitations, and Assumptions .............................................................. 16

  Definitions of Key Terms ............................................................................................. 17

  Summary ....................................................................................................................... 18

CHAPTER TWO: REVIEW OF RELATED LITERATURE ...................................................... 19

  History of Teacher Certification ...................................................................................... 19

  Basic Knowledge Testing and Accountability ............................................................. 28

  Test Bias ......................................................................................................................... 32

  Systemic Change .......................................................................................................... 37

CHAPTER THREE: RESEARCH DESIGN AND METHODOLOGY ..................................... 42

  Problem Statement ....................................................................................................... 43

  Research Purpose ....................................................................................................... 43

  Research Questions and Null Hypotheses .................................................................. 43
Methodology .............................................................................................................45
Research Design ........................................................................................................45
Independent Variables ............................................................................................46
Dependent Variables ...............................................................................................46
Study Group .............................................................................................................46
Data Collection .........................................................................................................47
Confidentiality and Data Anonymity Assurances ....................................................48
Instrumentation ........................................................................................................49
Data Analysis ............................................................................................................49
Summary ....................................................................................................................51

CHAPTER FOUR: PRESENTATION AND ANALYSIS OF DATA ................................53
Review of Problem and Purpose of the Study ........................................................53
Instrumentation Review ..........................................................................................54
Organization of Data Analysis ................................................................................54
Population of the Study Group ..............................................................................55
Demographic Characteristics of the Study Group ...............................................55
Research Questions and Null Hypotheses ..............................................................61
Analysis of Data .......................................................................................................63
  Research Question 1 ..............................................................................................63
  Research Question 2 ..............................................................................................76
  Research Question 3 ..............................................................................................81
  Research Question 4 ..............................................................................................83
  Research Question 5 ..............................................................................................92
Research Question 6 .................................................................95

Summary ..................................................................................100

CHAPTER FIVE: FINDINGS, CONCLUSIONS, IMPLICATIONS, AND
RECOMMENDATIONS ......................................................................103

Summary of the Study ...............................................................103

Problem Statement .....................................................................104

Purpose of the Study .................................................................104

Research Questions and Null Hypotheses .................................105

Review of Related Literature .....................................................106

Study Group ...............................................................................107

Findings ....................................................................................107

Research Question 1 ...............................................................108

Research Question 2 ...............................................................109

Research Question 3 ...............................................................109

Research Question 4 ...............................................................110

Research Question 5 ...............................................................113

Research Question 6 ...............................................................114

Conclusions ..............................................................................116

Research Questions 1, 2, and 3 .................................................117

Research Question 4 ...............................................................117

Research Question 5 ...............................................................119

Research Question 6 ...............................................................120
Implications ..................................................................................................................121

MoGEA and Test Bias ....................................................................................................121

Predictor Variables for MoGEA Success ....................................................................122

Recommendations for Future Research ......................................................................125

Summary .......................................................................................................................127

REFERENCES ...............................................................................................................129

APPENDIX A ................................................................................................................133

VITA ..............................................................................................................................134
LIST OF FIGURES

Figure .................................................................................................................. Page

1. Histogram of Total Population Total Credit Hours ........................................55
2. Histogram of Total Population GPAs .............................................................56
3. Histogram of Total Population Composite ACT Scores ...............................58
4. Histogram of Total Population MoGEA English & Language Arts Scores ....61
5. Histogram of Total Population MoGEA Math Scores .................................62
6. Histogram of Total Population MoGEA Science Scores ...............................62
7. Histogram of Total Population MoGEA Social Studies Scores ....................63
8. Histogram of Total Population MoGEA Writing Scores ...............................63
9. Histogram of Moderately Selective MoGEA English and Language Arts Scores ...............................................................65
10. Histogram of Moderately Selective MoGEA Math Scores ............................66
11. Histogram of Moderately Selective MoGEA Science Scores .........................66
12. Histogram of Moderately Selective MoGEA Social Studies Scores ...............67
13. Histogram of Moderately Selective MoGEA Writing Scores .......................67
14. Histogram of Selective MoGEA English and Language Arts Scores ............68
15. Histogram of Selective MoGEA Math Scores ................................................69
16. Histogram of Selective MoGEA Science Scores ..........................................69
17. Histogram of Selective MoGEA Social Studies Scores ...............................70
18. Histogram of Selective MoGEA Writing Scores ..........................................70
19. Q-Q Plot of Total Population MoGEA English and Language Arts Scores ....74
20. Q-Q Plot of Total Population MoGEA Math Scores ....................................74
21. Q-Q Plot of Total Population MoGEA Science Scores..................................75

22. Q-Q Plot of Total Population MoGEA Social Studies Scores..........................75

23. Q-Q Plot of Total Population MoGEA Writing Scores...................................76
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Research Questions and Null Hypotheses</td>
<td>15</td>
</tr>
<tr>
<td>2. Summary of Statistical Analyses by Research Question</td>
<td>51</td>
</tr>
<tr>
<td>3. Institutional Selectivity of Participating EPPs</td>
<td>56</td>
</tr>
<tr>
<td>4. Ethnicity Breakdown of Teacher Candidates</td>
<td>56</td>
</tr>
<tr>
<td>5. Gender Breakdown of Teacher Candidates</td>
<td>57</td>
</tr>
<tr>
<td>6. Total Credit Hours of Teacher Candidates</td>
<td>57</td>
</tr>
<tr>
<td>7. GPAs of Teacher Candidates</td>
<td>58</td>
</tr>
<tr>
<td>8. Financial Aid Breakdown of Teacher Candidates</td>
<td>59</td>
</tr>
<tr>
<td>9. Class Level Breakdown of Teacher Candidates</td>
<td>60</td>
</tr>
<tr>
<td>10. Composite ACT Scores of Teacher Candidates</td>
<td>60</td>
</tr>
<tr>
<td>11. Summary Statistics of Total Population MoGEA Subtest Scores</td>
<td>64</td>
</tr>
<tr>
<td>13. Kolmogorov-Smirnov Normality Test by MoGEA Subtest</td>
<td>76</td>
</tr>
<tr>
<td>14. MoGEA Subtest Mean and Median Scores by Decile</td>
<td>82</td>
</tr>
<tr>
<td>15. ANOVA Results for ACT and MoGEA Subtest Scores</td>
<td>84</td>
</tr>
<tr>
<td>16. ANOVA Results for Ethnicity and MoGEA Subtest Scores</td>
<td>85</td>
</tr>
<tr>
<td>17. Post Hoc (Tukey) Results for Ethnicity and Social Studies Subtest</td>
<td>87</td>
</tr>
<tr>
<td>18. ANOVA Results for Class Level and MoGEA Subtest Scores</td>
<td>87</td>
</tr>
<tr>
<td>19. Post Hoc (Tukey) Results for Class Level and Math Subtest</td>
<td>88</td>
</tr>
<tr>
<td>20. Independent-Samples t-Test Results for Ethnicity</td>
<td>89</td>
</tr>
<tr>
<td>21. Independent-Samples t-Test Results for Gender</td>
<td>90</td>
</tr>
</tbody>
</table>
22. Independent-Samples $t$-Test Results for Financial Aid .................................................91
23. MoGEA Subtest Passing Rates of Ethnic Groups at Various Cut Scores ..............93
24. PPMCC Results for MoGEA Subtest Scores and Independent Variables.............96
25. MANOVA Results for Independent Variables ..........................................................98
26. Discriminant Functional Analysis Results for Ethnicity, Gender, and Financial Aid.........................................................................................................................99
ABSTRACT

The purpose of this quantitative study was to recommend statewide cut score norms for the Missouri General Education Assessment (MoGEA) to the executive board of the Missouri Association of Colleges for Teacher Education for the purpose of informing its representatives to the Missouri Advisory Council of Certification for Educators. Additionally, this study aimed to identify demographic characteristics of educator preparation program (EPP) candidates that might serve as predictors of success on the MoGEA. If predictor variables are determined, EPP administrators can identify which of their education students will most likely require additional support and preparation for the MoGEA. This information will also make Missouri’s Department of Elementary and Secondary Education (DESE) and Missouri’s State Board of Education aware of which EPP candidate demographics are likely to be denied EPP admission.

The statistical findings of the study revealed significant differences in how teacher candidates from different demographic groups performed on the MoGEA subtests during the 2013-2014 academic year. Through various statistical analyses conducted in response to the study’s six research questions, test biases against specific populations were revealed and predictor variables emerged. Test biases against teacher candidates from the Black, Non-Hispanic and Hispanic/Latino populations and also against female teacher candidates were of primary concern to the researcher. Composite ACT score
emerged as the strongest predictor variable of success on the MoGEA subtests, but other demographic characteristics could also be used by EPP to identify candidates who will potentially struggle on the MoGEA, including ethnicity, gender, total credit hours, cumulative GPA, and whether or not a student receives financial aid.

The researcher ultimately concludes the MoGEA should not be used as a qualifying examination for teacher candidacy and that alternative means of determining whether candidates are qualified be explored. The researcher also recommends the MoGEA undergo additional review, development, and field-testing before statewide cut score norms are set by DESE due to the test biases revealed through statistical analyses of the MoGEA scores collected from 1037 teacher candidates from ten different EPPs throughout the state. The researcher provides EPP personnel with suggestions for the identification and support of teacher candidates whose personal demographic characteristics indicate potential difficulty obtaining passing scores on this examination. The researcher also recommends continued collection and analysis of MoGEA scores from demographic groups underrepresented in this study.
CHAPTER ONE

INTRODUCTION TO THE STUDY

The ultimate goal of this study was to analyze teacher candidate scores on the five subtests of the Missouri General Education Assessment (MoGEA) for the 2013-2014 school year and recommend qualifying cut scores for each subtest. This study aimed to consolidate MoGEA subtest scores of teacher candidates throughout the state of Missouri to gain a broader view of teacher candidate performance on this assessment in its first year of use. Additionally, it was a goal of this study to determine if any demographic characteristics could be used to predict student success on the MoGEA subtests, based on student demographic data provided by each educator preparation program (EPP) that self-selected to be a part of this research project. Finally, through data analysis and the examination of past research regarding basic knowledge tests used for teacher certification, the researcher aimed to recommend cut scores for each subtest to be used in future years by EPPs statewide. To better understand the larger issues associated with this study, the researcher examined the following themes: the history of teacher certification throughout the United States, the use of basic knowledge tests as a means of accountability by teacher preparation programs, and the potential biases involved in the use of basic knowledge tests. Investigation of these themes greatly influenced the researcher’s perspective throughout the entire study and was of chief importance when attempting to determine statewide cut scores.

This first chapter serves to introduce the reader to the study. First, the major themes involved in the study will be introduced, including a brief examination of the history of teacher certification in the United States. As part of this historical account, the
role basic knowledge tests have played in measuring accountability throughout the educational system will be reviewed. Additionally, the potential bias basic knowledge tests required for teacher licensure have against specific teacher candidate populations will be acknowledged. Second, the conceptual underpinning for this study will be explained. Next, the problem in this study and the purpose of the research will be described. The study’s research questions and null hypotheses will then be identified in narrative form and displayed in Table 1. A discussion of the limitations, delimitations, and assumptions of the project follow the presentation of the research questions and null hypotheses. Finally, the key terms involved in the study will be defined.

**Background**

The topics of requirements for teacher certification and accountability for teacher quality are not new to the educational discourse. They have, however, come to dominate the current dialogue regarding the educational system in the United States (Feuerstein, 2014). Both educational professionals and state and federal legislators now seem well-versed in deciding what teachers need to know in order to effectively teach their students, how teachers should demonstrate this knowledge, and whom should be held responsible when students do not show signs of growth or mastery. Unfortunately, the inability of these two parties to agree on these decisions has created great tension related to the state of public education throughout the US. This tension is not a recent concern: “Teacher education in the United States has long had a complicated relationship with structures of governmental power and control. Local, state, and federal roles in teacher training and certification have been inconsistent throughout the nation’s history” (Tobin, 2012, p. 485). From the early 20th century when teacher certification was determined by local
district politicians or clergymen through the 1980s and the movement toward teacher certification at the national level, control over the credentialing of the nation’s teacher candidates has gradually become more centralized (Bales, 2006; Tobin, 2012). The power to certify teachers passed from local authorities to state-governed educational bureaucracies (Tyack & Cuban, 1997), and, while individual states still have the authority to grant certification, the national government now has a profound influence over the standards to which states must hold their teacher candidates (Tobin, 2012).

The shift of power from local districts to state and national governments resulted from the fluctuating imbalance of teacher supply and demand experienced throughout the 1900s (Angus, 2001). Historical events, such as the Great Depression, World Word II (WWII), and the Cold War repeatedly tilted this balance back and forth, and state and national governments responded by increasing or decreasing the certification requirements for educators. When teachers were plentiful but jobs were scarce, it became more difficult to earn a teaching credential. Likewise, when the United States experienced teacher shortages, teacher candidates more easily obtained certification (Angus, 2001; Darling-Hammond, 1999; Tobin, 2012).

Similar to the manner in which historical events affected the certification process for teacher candidates, federal legislative actions also played a significant role in the centralization of teacher certification (Tobin, 2012). The National Defense Education Act of 1958, the Elementary and Secondary Act of 1965, Goals 2000: The Educate America Act of 1994, and the No Child Left Behind Act of 2001 (NCLB) each served to further centralize the power of teacher certification (Bales, 2006; Cibulka, 1990). The impact of NCLB in particular is still felt by states today as they struggle to address its long-term
guidelines for teacher preparation and student achievement: “The inclusion of a reported link between teacher quality and student achievement is novel in the NCLB law, and efforts to quantify this effect presently dominate the national education policy agenda” (Tobin, 2012, p. 490). Corresponding with these acts of federal legislation were the creation and implementation of different means of holding educators, administrators, and the entire educational system accountable for the success of the nation’s students (Feuerstein, 2011). Unfortunately, as Feuerstein (2011) explains, the various legislative actions taken to hold members of the educational community accountable for student progress “have resulted in an overly rigid system of accountability focused on testing and sanctions that may actually work against improved school performance by narrowing debate, dehumanizing education, and distracting teachers from their efforts to engage students in authentic educational experiences” (p. 7).

Evidence of these pitfalls and other problematic results of federal attempts to control teacher certification and hold educators accountable, are the consequences of the passage of NCLB and Race to the Top (RttT). Although the federal government passed NCLB as a further attempt to centralize the certification of the nation’s teachers, it is interesting to note that even more than a decade prior to NCLB almost all U.S. states had common requirements for teacher licensure:

In all but a few [states], teachers had to have proof of an average of 30 undergraduate or graduate credits from a state-approved teacher education institution, 12 to 15 weeks of student teaching experience, a major or minor in their subject area, and high-enough test scores in basic skills as well as subject matter and/or pedagogical techniques. (Tobin, 2012, p. 486)
While states still separately controlled teacher licensure within their own borders, they already had a similar idea of what teachers should know in order to perform effectively in the classroom and the training teachers need in order to acquire this knowledge. In this case, without and prior to federal involvement, states already had common measures of teacher accountability.

President Barack Obama and Secretary of Education Anne Duncan announced the most recent act of federal legislation affecting teacher certification, Race to the Top, in the summer of 2009. Unlike NCLB’s focus on staffing classrooms with effective teachers, however, RttT emphasizes student achievement and links elementary and secondary students’ scores on standardized tests to the EPP through which their teachers became certified. Race to the Top, therefore, was the first act of federal legislation containing stipulations specifically aimed at holding EPPs accountable for teacher effectiveness, as measured by student test scores (Henry, Kershaw, Zulli & Smith, 2012; Lewis & Young, 2013). RttT, like NCLB fueled the desire by legislators for increased testing of both K-12 students and teacher candidates as a measure of accountability.

Additionally, while NCLB and RttT dictate specific guidelines by which school districts and states must abide in order to receive federal funding, there is much latitude for interpretation of these guidelines (Selwyn, 2007; Superfine, 2005). Of particular significance is the ambiguous mandate for all core teachers to be “highly qualified”, which, according to the law requires teachers to be “fully certified” (Tobin, 2012; U.S. Department of Education, 2008). It is this requirement of full certification that each state is presently left to define for itself. Many states have responded to NCLB’s requirement to ensure educators are “highly qualified” and “fully certified” by increasing the amount
of high-stakes standardized tests required for admission into (EPPs) and for teacher certification.

While these tests of general knowledge, curricular content, and pedagogy are designed to assess whether a teacher candidate is competent and qualified, many education professionals argue such tests do not and cannot measure candidates’ dispositions or if they are able to apply acquired pedagogical skills in classroom settings, both of which are critical aspects considered very important for successful teaching. As Goodman, Arbona, and de Rameriz (2008) stated, “effective teachers must not only know what to do; they must also be willing and able to do it” (p. 24). Not only is the research mixed regarding whether these tests of basic and content area knowledge are effective tools for measuring teacher quality, but there is also much evidence supporting the claim that these tests are actually biased against minority populations and prevent many would-be minority teachers from becoming licensed educators (Gitomer, Brown, & Bonett, 2011; Goodman, Arbona, & de Rameriz, 2008; Wakefield, 2008), an assertion that will be further explored in the second chapter of this study.

Despite the position held by many researchers that standardized tests of basic knowledge and content knowledge cannot appropriately measure a teacher candidate’s potential effectiveness in the classroom, almost every state—46 as of 2006—now require such tests for teacher licensure (Goldhaber, 2007). The state of Missouri is among these and is of primary concern to this study. Missouri’s Department of Elementary and Secondary Education (DESE) currently offers six different routes to teacher certification: traditional, alternative or innovative, temporary authorization, out-of-state certified, American Board of Certification for Teacher Excellence (ABCTE), and doctoral. Of
these six, only the out-of-state certified route does not require passing at least one basic knowledge or content area examination (Missouri, 2014). Beginning in the late 1980s, Missouri’s EPPs began using the College Basic Academic Subjects Examination (C-BASE) as an assessment of teacher candidates’ basic knowledge. In order to be accepted into an EPP, teacher candidates had to pass each of the subtests on the exam, obtaining scores at or above the cut scores mandated by Missouri’s State Board of Education (Assessment, 2012).

While national and state legislators and educational professionals do not always share a common opinion regarding the use of basic knowledge exams as a requirement for teacher certification (Goldhaber, 2007; Tobin, 2012), when such exams have been in place for an extended period of time, EPP administrators and educators are able to create and implement specific programs and strategies for test-taking as part of the EPP curriculum and experience. In Missouri, for example, because the C-BASE had been used since the late 1980s, EPP administrators were familiar with the exam and some had even developed test preparation strategies for their teacher candidates. Wall (2008) wrote specifically of C-BASE preparation strategies used by one EPP in Missouri and the positive effect they had on students’ success in earning qualifying C-BASE scores.

Given years to create and implement supporting measures, EPP administrators were able to better prepare their students for success on state-mandated tests, utilizing strategies such as thoughtful curriculum alignment and the implementation of test-preparation programs, thus demonstrating their ability to prepare teacher candidates to meet state and federal requirements for licensure. If a change is made to the licensure exam for which EPP administrators have prepared their students, student performance
and success rates can be expected to change. Such is the case in the state of Missouri: On August 23, 2013, a memo from Missouri’s Director of Educator Preparation informed EPP students, faculty and staff, and administrators that the state would no longer be using the C-BASE as the requisite assessment of general education skills and knowledge. Instead of the C-BASE, effective September 1, 2013, DESE would implement a new qualifying assessment to determine whether Missouri’s teacher candidates possess the basic knowledge necessary to progress further in the process toward becoming state-certified educators (G. Hairston, personal communication, August 21, 2013).

In September of 2013, the Missouri General Education Assessment was first used as this new qualifying exam for EPPs throughout the state of Missouri (G. Hairston, personal communication, April 16, 2014). To be admitted into one’s institution’s EPP, a student would need to achieve a minimum cut score on each of the MoGEA’s five sections: English language arts, writing, math, science, and social studies (MEGA, 2013). Because the MoGEA was a newly created exam and there were no prior scores to be used as a basis for setting minimum required scores across the state, DESE permitted each individual EPP to determine its own cut scores during the 2013-2014 academic year (Pearson, 2014). Following the 2013-2014 school year, a committee of Missouri educators planned to convene and make a formal recommendation to DESE regarding statewide cut scores. The final decision regarding qualifying scores for admission into EPPs would rest in the hands of Missouri’s BOE and would be made in June of 2015. Of chief importance to this study is whether this shift from the C-BASE to the MoGEA will affect the success of candidates for admission to EPPs throughout the state of Missouri.
Conceptual Underpinnings

The conceptual underpinning guiding this study is the process of systemic change, about which researchers including Banathy, Reigeluth, Roberto, and Jenlink wrote extensively. Systemic change “recognizes the interrelationships and interdependencies among the parts of the educational system, with the consequence that desired changes in one part of the system must be accompanied by changes in other parts that are necessary to support those desired changes” (Jenlink, Reigeluth, Carr, & Nelson, 1998, p. 219). This is an appropriate guiding framework for this study because the shift from the use of the C-BASE to the MoGEA as the qualifying assessment of basic knowledge for Missouri’s EPP candidates represents a statewide policy change that affects the entire system of teacher education in Missouri. The process of systemic change acknowledges the effect a change in one component of a system has on all other components. As summarized by Reigeluth (1992):

Systemic change is comprehensive. It recognizes that a fundamental change in one aspect of a system requires fundamental changes in other aspects in order for it to be successful. In education, it must pervade all levels of the system: classroom, building, district, community, state government, and federal government. (p. 3)

In the particular case of this study, DESE took what Squire and Reigeluth (2000) referred to as a “statewide policy approach” to the process of systemic change within an educational setting. Statewide approaches often involve the coordination of statewide tests and requirements for teacher certification, and “usually, a fairly small set of people representing a narrow portion of educational interests is involved in [the] legislative
process” (p. 147). According to Missouri’s Director of Education Preparation, Dr. Gale Hairston, the decision to change from the C-BASE to a different assessment was made by the Missouri’s State Board of Education (BOE), the Commissioner of Education, and Dr. Hairston’s office (personal communication, September 16, 2014), which aligns with Squire and Reigeluth’s definition of a “statewide policy approach” to change. Further connections between the change from the C-BASE to the MoGEA and the work of these two authors can be made to their description of the involvement, or the lack thereof, of various stakeholders in the statewide policy change process:

This approach to educational reform…involves experts at the state level informing system members on how educational systems should operate. Teachers, administrators, and students do not have much voice in the goal-setting or decision-making processes; rather, they are expected to adopt the policies generated by the statewide systemic reformers. (2000, p. 147)

Squire and Reigeluth (2000) acknowledged that key stakeholders are ignored when educational systems take this type of approach to systemic change. When certain stakeholder groups are ignored in the decision making process leading to a change in policy, and these same groups are also heavily affected by the ramifications of the decision, the process of change will oftentimes be particularly challenging.

In conjunction with the research of Squire and Reigeluth, Banathy’s work further expounds upon the importance of involving all stakeholders in the change process. In his writing on the conceptual framework of systemic change, Banathy identified six specific process values that serve as the foundation for the ways members of the educational community think about the design of the system: commitment to participation,
commitment to idealized design, design is continuous, nurturing human values and human quality, design community, and design conversation (Joseph, Jenlink, Reigeluth, Carr-Chelman, & Nelson, 2002). Joseph et al. (2002) defined these six process values as “the intrinsic qualities that ground and guide the collective set of beliefs we share and uphold as we travel through a journey in creating a better educational system for our communities” (p. 380). Educators’ values and beliefs certainly guide their decision making, and, as Joseph et al. explained, systemic change will only be successful when all stakeholders share such beliefs and values.

The first of these process values, commitment to participation, asserts the importance of involving stakeholders in all aspects of the change process (Joseph et al., 2002). Additionally, Joseph et al. (2002) also explain that stakeholders should not have change thrust upon them or designed by people other than themselves; those responsible for implementing the change should be the ones responsible for designing the process by which the change will occur.

The second process value, commitment to idealized design, calls those responsible for enacting the process of systemic change to strive toward the attainment of their ideal outcome, regardless of actual parameters, logistics, and limitations. Joseph et al. (2002) explained it is important to work toward the ideal system even though it might not be a realistic goal because “it is the dreams and aspirations we share and act upon that will bring us closer to the ideal educational systems we seek to create for our future, for our society, and for our children” (p. 382). In conjunction with a commitment to idealized design, Banathy’s third process value, design should be continuous, prescribes constant revision and redesign of the change process being undertaken. This is necessary because
any process of change is always evolving and new opportunities, ideas, and obstacles will organically present themselves as work is done (Joseph et al., 2002).

The fourth systemic change process value, nurturing human values and human quality, contains seven components: caring for children and their future, respect, social responsibility, social justice, community, creativity, and collaboration. Together, these seven components acknowledge and consider the various stakeholders and collectively recognize human beings as the top priority in the change process. These components also serve to provide the overall process with a moral compass fixed on the ideal outcome for all who are involved with and affected by the experienced change (Joseph et al., 2002).

The fifth and sixth process values identified by Banathy, design community and design conversation, describe the importance of all stakeholders sharing a common goal and committing to its attainment while communicating with one another in ways that “ensure inclusive participation of those who are beneficiaries and users of the system being created” (Joseph et al., 2002, p. 385). In Chapter Two of this study, each of these six process values will be further examined as it relates to the reviewed literature, the study’s problem, and the research purpose.

**Problem Statement**

At the time of this study’s proposal, no statewide qualifying scores existed for the Missouri General Education Assessment. In order for DESE to establish examination cut scores based on the input of Missouri’s EPP personnel and inform the final decision made by the BOE, it was critical that statewide results of the 2013-2014 MoGEA be collected and analyzed to ensure a more representative population of test scores be used to establish valid and reliable cut scores for use in 2015.
Research Purpose

The purpose of this quantitative study was to recommend statewide cut score norms for the MoGEA to the executive board of the Missouri Association of Colleges for Teacher Education (MACTE) for the purpose of informing its representatives to the Missouri Advisory Council of Certification for Educators (MACCE). Additionally, this study aimed to identify demographic characteristics of EPP candidates that might serve as predictors of success on the MoGEA. If predictor variables are determined, EPP administrators can identify which of their education students will most likely require additional support and preparation for the MoGEA. This information will also make DESE and the BOE aware of which EPP candidate demographics are likely to be denied EPP admission.

Research Questions and Null Hypotheses

This study was designed to address the following research questions and their accompanying null hypotheses:

1. What are the descriptive summary statistics for each MoGEA subtest for the total population of the study group and when broken out by institutional selectivity?

2. Do statewide scores on the MoGEA subtests for the 2013-2014 academic year represent a normal distribution?

   $H_0$: Statewide scores on the MoGEA for the 2013-2014 academic year do not represent a normal distribution.

3. What are the mean and median scores for each decile within each MoGEA subtest?
4. When composite ACT score is used as a covariate, is there a difference in MoGEA mean subtest scores for the following demographic variables?
   a. Ethnic origin
   b. Gender
   c. Total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained
   d. Cumulative grade point average at the time the student’s highest reported score on the MoGEA was obtained
   e. Classification of the candidate as receiving financial aid or not receiving financial aid

   \[ H_0^2: \text{There is no difference in MoGEA mean subtest scores between demographic variables.} \]

5. How will establishing cut scores at the lowest score within each decile affect teacher candidate admission for populations that have historically been shown to be negatively affected by test bias?

   \[ H_0^3: \text{Establishing cut scores at the lowest score within each decile has no effect on teacher candidate admission for populations that have historically been shown to be negatively affected by test bias.} \]

6. Is it possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study?

   \[ H_0^4: \text{It is not possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study.} \]
Table 1 displays the study’s six research questions and corresponding null hypotheses.

Table 1

<table>
<thead>
<tr>
<th>Research Questions &amp; Null Hypotheses</th>
<th>Research Question</th>
<th>Null Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>What are the descriptive summary statistics for each MoGEA subtest for the total population of the study group and when broken out by institutional selectivity?</td>
<td>H₀₁: Statewide scored on the Missouri General Education Assessment for the 2013-2014 academic year do not represent a normal distribution.</td>
</tr>
<tr>
<td>2.</td>
<td>Do statewide scores on the MoGEA subtests for the 2013-2014 academic year represent a normal distribution?</td>
<td>H₀₂: There is no difference in MoGEA mean subtest scores between demographic variables.</td>
</tr>
<tr>
<td>3.</td>
<td>What are the mean and median scores for each decile within each MoGEA subtest?</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>When composite ACT score is used as a covariate, is there a difference in MoGEA mean subtest scores for the following demographic variables?</td>
<td>H₀₃: Establishing cut scores at the lowest score within each decile has no effect on teacher candidate admission for populations that have historically been shown to be negatively affected by test bias.</td>
</tr>
<tr>
<td></td>
<td>a. Ethnic origin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b. Gender</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c. Total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d. Cumulative grade point average at the time the student’s highest reported score on the MoGEA was obtained</td>
<td></td>
</tr>
<tr>
<td></td>
<td>e. Classification of the candidate as receiving financial aid or not receiving financial aid</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>How will establishing cut scores at the lowest score within each decile affect teacher candidate admission for populations that have historically been shown to be negatively affected by test bias?</td>
<td>H₀₄: It is not possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study.</td>
</tr>
<tr>
<td>6.</td>
<td>Is it possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study?</td>
<td>H₀₄: It is not possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study.</td>
</tr>
</tbody>
</table>
Limitations, Delimitations, and Assumptions

As a natural product of the design and execution of this study, limitations and delimitations were foreseen by the researcher. Regarding the data collection process of this study, it was not possible to conduct simple random sampling of students’ MoGEA scores, limiting the generalizability of the study’s findings (Nenty, 2009). Additionally, while numerous attempts were made to collect data from all EPPs in the state of Missouri, individual EPP involvement in the study was voluntary; participation was self-selected by each EPP in the state, which narrowed the scope of the study and afforded EPP personnel the opportunity to opt-out of the research. Also, during the processes of data collection and analysis it was necessary to limit the focus of the study to the specific demographic variables detailed in the research questions guiding this project. A further limitation regarding the demographic variables utilized in this study was the lack of availability of this information for all subjects in the study group, which affected the sample sizes associated with each independent variable in the study’s fifth research question. Finally, in order to make a timely recommendation to DESE, it was not possible to delay data analysis until participation by all EPPs could be guaranteed, thereby affecting this study’s sample size. Further limitations that became apparent during the data analysis portion of this study will be discussed in Chapter 5.

In conjunction with accounting for delimitations and limitations of this study, the researcher also acknowledges various assumptions made throughout the course of the research. Because the researcher had been a classroom teacher for over ten years at the time of the study and was trained and certified in the state of Missouri, effort was made to remain unbiased in regard to issues and events associated with this study. Secondly, it
was assumed that students performed to the best of their ability when taking the MoGEA, giving maximal effort on every subtest. Finally, because the data collection process required EPP personnel to consolidate student scores and demographic information from various resources into one summarizing spreadsheet, it was assumed this was done accurately and that scores for all students who had taken the MoGEA at each participating institution were included in the contributed data.

**Definitions of Key Terms**

Because this study required the use of specific vocabulary, the following terms are defined for the reader in order to clarify terminology.

*ACT.* The ACT Assessment is a widely used college entrance examination that assesses high school students’ general educational development and their ability to complete college-level work (Wall, 2008).

*C-BASE.* The C-BASE (College Basic Academic Subjects Examination) is a criterion-referenced achievement examination that qualifies individuals for EPPs and tests general academic knowledge and skills (Assessment, 2012; Wall, 2008).

*Cut score.* A cut score is the minimum score required to pass an examination.

*DESE.* DESE (Missouri Department of Elementary and Secondary Education) is the governing body of Missouri public education.

*Educator preparation program.* Educator preparation programs are state-board approved, state and/or nationally accredited institutions of higher education responsible for developing teacher candidate knowledge and providing them with necessary skills and experiences to earn licensure at the completion of particular programs of study. This term is used interchangeable with *teacher preparation program* (TPP).
High-stakes tests. In this study, high-stakes tests are standardized examinations that have “serious consequences” for teacher candidates and teacher education programs, including licensure and accreditation decisions for candidates and programs, respectively (Jones, Jones, & Hargrove, 2003, p. 3; Wall, 2008).

MoGEA. MoGEA (Missouri General Education Assessment) is the test of basic knowledge required for admission into EPPs in the state of Missouri. It consists of five subtests: English language arts, writing, math, science, and social studies. Students must achieve passing scores on all subtests in order to be admitted into their institution’s EPP (Pearson, 2014).

Summary

Tobin (2012) skillfully summarized one of the major themes of this study: “Issues of teacher training and certification have always revolved around the central question of who should control these processes: education professionals, government, or public agencies” (p. 485). Throughout history, power over teacher certification has gradually become more centralized, shifting away from individual EPPs and toward state and federal government agencies. The decision to change from the C-BASE to the MoGEA as the qualifying basic knowledge assessment used by all EPPs throughout the state is representative of this centralization of control and affects all of Missouri’s EPP administrators, faculty, and students. As will be further explored in the second chapter of this study, attempts at change to any part of a system have far-reaching consequences and greatly affect all stakeholders within the system. The educational system, both within the state of Missouri and at the national level, is no exception. What might at first seem to be a simple change in policy could potentially be much more complicated.
CHAPTER TWO

REVIEW OF RELATED LITERATURE

As was introduced in the first chapter of this study, teacher certification in the United States and the specific requirements for licensure of teacher candidates historically depended upon the relationship between teacher supply and demand and was heavily influenced by key events and federal legislative actions. For the purposes of this study, the review of literature will first further expand upon the history of teacher certification in the United States by highlighting historical events and legislative actions that influenced licensure requirements. Second, and of particular importance to this study’s research questions, the use of basic knowledge tests as a measure of accountability for teacher quality and certification will be explored. As part of this topic, research regarding the history of basic knowledge tests for teachers, the advantages and disadvantages of requiring teacher candidates to pass such tests, and their effectiveness as predictors of teacher success in the classroom will be discussed. Third, the issue of test bias as it relates to minority educator preparation program (EPP) candidates will be addressed. These three topics will then be examined through the lens of systemic change, the process that underpins the conceptual framework for this study.

History of Teacher Certification

The history of teacher certification in the United States can be traced back to the early 1800s, when the general belief regarding what was required for an individual to become a teacher can be summarized by the following quote from an historical manifesto from 1875: “If a teacher has a love of knowledge, and has a strong and quick feeling for childhood, a few simple and easily taught rules, devices, and a few dozen lessons...are
enough for the rank and file” (as cited in Blackford, Olmstead, & Stegman, 2012). During the 1830s, however, schools began the transition from privately funded educational institutions to free public schools. Throughout the country, children of rural populations received their education in one-room schoolhouses while urban schoolchildren attended multi-classroom schools serving much larger populations. This dichotomy in school structure and population and its effect on the concept of free public education fueled the first debates regarding teacher training and preparation as stakeholders called for equal qualifications for teachers in both settings (Blackford et al., 2012). Up until this time, the certification of teachers was decentralized and had been solely of local concern. Urban and rural schools trained and licensed teachers differently, as politicians or clergymen in a school’s vicinity were responsible for awarding licenses (Tobin, 2012).

Power steadily began to shift in 1843, when the state of New York enacted the first piece of legislation that put the authority over teacher licensure in the hands of the state rather than local leaders. Other states followed its lead over the course of the next fifty years. By the end of the 19th century, most states granted certification based on teacher examinations that assessed basic knowledge and general understanding of U.S. history, geography, spelling, and grammar (Blackford et al., 2012). As the 19th century closed and the 20th century began, free public education was quickly expanding in response to the nation’s growing population. As Hess (2009) explained, the teaching profession, specifically the recruitment and certification of teachers, had to respond to this growth. Educated women were recruited to become teachers as most other professions were not available to female jobseekers, and the teaching profession “assumed little in the way of a professional knowledge base, and teacher preparation and
development were designed accordingly” (p. 450). Certification was primarily awarded based on the passing of basic knowledge tests as part of the successful completion of a college or university EPP (Tobin, 2012).

As the 1930s approached, control over teacher certification continued to become more centralized as formal educational bureaucracies working for state governments began determining the requirements for teacher licensure. Members of state governments and these bureaucracies believed an increase in the amount of teacher training would increase the level of teacher quality and the requirements for certification subsequently began to increase. Further continuing this trend toward stricter state-mandated requirements was a shift in teacher supply and demand during the Great Depression. In order to counterbalance the nation’s oversupply of teachers, certification requirements increased, serving to lessen the number of qualified jobseekers. State governments continued to gain even more control over teacher licensure, and by 1937, 41 states had systems in place where teachers were awarded certificates solely by the state (Tobin, 2012).

The next significant change in requirements for teacher certification was the result of a dramatic shortage in teachers following the United States’ entry into World War II (WWII). As women left the teaching profession to take the jobs of men serving in the armed forces, teaching vacancies rapidly increased throughout the country. School administrators and state legislators needed to respond to the teacher shortage and did so by lessening the requirements for teacher licensure. Fewer examinations were required and the quality of teachers’ credentials was no longer a chief concern; the nation simply needed teachers (Blackford et al., 2012; Tobin, 2012). The increase in the number of
emergency certificates issued to teacher candidates from 1940 until the end of the war serves as evidence of this need. In 1940, 2,305 emergency certificates were issued. By the end of the war, the number had skyrocketed to 108,932 (Tobin, 2012). This decrease in teacher certification requirements during and immediately following WWII is believed to have caused an overall drop in teacher quality and, as a result, student learning and achievement (Blackford et al., 2012).

In 1957, when the Soviet Union launched Sputnik into outer space, that the quality of the nation’s teachers and the programs that prepared them once again returned to the forefront of the nation’s concerns: “The Soviets beating Americans into space was a wake-up call that sparked officials to attack teacher preparatory schools” (Blackford et al., 2012, p. 7). Schools staffed with teachers who had been granted emergency certificates during WWII had not adequately prepared students in mathematics and the sciences, and the Soviet Union was perceived to be far ahead of the United States in these subject areas. This perception served to further intensify this threat of Communism posed by the Soviet Union. In response, the U.S. passed the National Defense Education Act of 1958, which provided funds for math and science education nationwide in hopes of raising a new generation of scientists capable of competing with the Soviet Union in the arms and technology race (Tobin, 2012). While this bill did not formally dictate teacher certification requirements, it did mark the first time the national government became officially involved in education policymaking (Bales, 2006).

Coinciding with a second teacher shortage as a result of the baby boom in the 1960s, a second act of national legislation, the Elementary and Secondary Education Act of 1965, served to further centralize control over education by creating the National
Assessment of Educational Progress (Tobin, 2012). This assessment allowed legislators to compare student achievement nationwide, expanding the role of the federal government in public education (Robelen, 2005). While this act, like the National Defense of Education Act, did not specifically mandate teacher certification requirements, the decades in which they were passed saw a renewed interest in improving teacher quality. States once again bolstered their certification requirements putting an even greater focus on content knowledge and teacher examinations (Blackford et al., 2012).

In the 1980s, control over teacher certification and preparation continued to become even more centralized. In their article “A Challenge for Leadership of Teacher Education in the Eighties,” Durkin and Smith (1981) expressed the following concern: People other than teacher educators are deciding the future of teacher education. Recent events show that the leadership for establishing requirements for certification and changing traditional preservice teacher education programs is shifting from teacher educators to legislators. Legislators are either passing laws specifically setting regulations, establishing standards, and prescribing programs for preservice teacher education or are directing state departments of education to do so. This phenomenon is a change from the previous latitude given to colleges or state departments of education. (p. 183)

Just two years later, the National Commission on Excellence in Education (NCEE) published its landmark report, *A Nation at Risk: The Imperative for Educational Reform*, effectively putting Durkin and Smith’s words into action. This report not only commented on the state of public education nationwide, it also marked the first time a
direct connection between teacher quality and student achievement was suggested, leading to further centralized attempts to control teacher licensure (Tobin, 2012). In fact, just ten years after the report was published, the number of states specifically requiring a standardized certification exam would reach 43 (Blackford et al., 2012). The NCEE report also advocated for a dramatic improvement in EPPS and that efforts be made to recruit teacher candidates of greater academic strength (Blackford et al., 2012; Tobin, 2012). An unintended yet serious consequence of *A Nation at Risk* was its effect on the general public perception of the nation’s teachers. The report “fostered the idea that the United States had fallen behind its peer countries in public education, in the process introducing the idea that American schools were ‘failing’” (National, 1983; Tobin, 2012, p. 487-8). While *A Nation at Risk* served to introduce teacher certification to the national dialogue regarding public education, it also “is perhaps the impetus for the antiteacher sentiment that characterizes much of school reform debate even today” (Tobin, 2012, p. 488).

Two additional acts of federal legislation, America 2000: An Education Strategy and Goals 2000: The Educate America Act, passed respectively in 1989 and 1994, also indirectly influenced EPPs and teacher certification. Neither act mandated specific requirements for teacher licensure, but they did represent legislative attempts to dramatically improve the state of public education nationwide. America 2000 is also credited with setting the stage for the dramatic increase in teachers obtaining licensure through alternative certification routes a decade later (Tobin, 2012). Similarly, Goals 2000 is acknowledged as “begin[ning] the accountability movement that has characterized recent education reforms” (Tobin, 2012, p. 489).
Historical events both within and outside the United States as well as federal legislature passed during the second half of the 20th century were the driving forces behind the shift in control over teacher certification and the specific requirements for licensure during the 1900s. The passing of the No Child Left Behind Act of 2001 (NCLB) and the announcement of Race to the Top in 2009, however, combined to affect the teaching profession and EPPs more heavily than any previous government actions. NCLB was the first act of federal legislation to specifically address the link between teacher quality and student achievement (Tobin, 2012). The law stipulates requirements schools must meet if they are to receive federal funding. One such stipulation requires teachers be highly qualified to teach the subject(s) to which they’ve been assigned:

“According to NCLB, teachers, to be highly qualified, must be certified by the state, have at least a bachelor’s degree, and pass basic skills and subject area tests” (Selwyn, 2007, p. 126). This requirement of state certification as part of federal law represented a dramatic shift of control over teacher licensure and placed most of the power in the hands of state governments. EPPs nationwide now had less authority over how their teacher candidates would be trained and evaluated. As Wakefield (2006) indicates, “NCLB allow[s] states to design or select their assessments for screening candidates and certifying or licensing teachers. The primary stipulation is that states must provide quality assurances in order to receive federal funding and avoid fines” (p. 81). EPPs nationwide were now required to include basic knowledge tests and content area examinations as part of teacher training and licensure (Tobin, 2012).
In conjunction with NCLB, Henry et al. (2012) explained how Race to the Top served to further hold EPPs accountable for teacher quality and, subsequently, student success:

Under RttT, states must develop and implement the capacity to reliably link student test scores to teachers to determine teacher effectiveness, and then link these measures of teacher effectiveness back to the programs that prepared them to teach. (p. 336)

Compounding the emphasis on the need for EPPs to prove the effectiveness of the teachers they train and certify, RttT pressures states to track student growth and improvement on standardized tests as measure for holding EPPs accountable. While RttT was not unique in linking student outputs to teacher inputs in this way, prior to this time, the technology and infrastructure necessary to effectively collect and analyze such large quantities of data simply did not exist (Henry et al., 2012).

Race to the Top’s requirement of the collection and analysis of this data served as the catalyst for the situation that is of primary concern to this study: Missouri’s shift from using the C-BASE as the qualifying basic knowledge test for admission to EPPs to the MoGEA. Missouri’s State Board of Education (BOE) desired a computer-based basic knowledge exam for its teacher candidates, which would significantly improve the state’s efforts to comply with federal data collection and reporting requirements, and the C-BASE was still a paper-and-pencil exam. Therefore, on September 19, 2012, Missouri’s Office of Administration released an official request for proposal seeking a new “entry assessment for teacher preparation” (Educator Assessments, 2012). In November of 2012, the BOE accepted a proposal from Pearson’s Evaluation Systems Group, and the
development process for the Missouri General Education Assessment began (G. Hairston, personal communication, September 18, 2014).

Following a mandatory 30-day waiting period after the proposal’s acceptance, Dr. Hairston and Missouri’s BOE held a meeting in December of 2012 to review Pearson’s proposed implementation plan. Thirty days later, in January of 2013, the BOE first communicated with Missouri’s EPPs to inform them of the upcoming change in qualifying exams. As part of the development process, Pearson’s Evaluation Systems Group worked from the state’s request for proposal to establish an initial framework for the exam. A bias review committee then reviewed the framework. Changes were made to the initial framework, including changes in test content to better reflect and account for the varying 42-hour general education requirements of Missouri’s colleges and universities; a new framework was then created (G. Hairston, personal communication, September 18, 2014).

After the creation of a second draft of the exam framework, test items for the writing portion of the MoGEA were created and field-tested during the summer of 2013 in order to ensure their reliability. In contrast to the prompts used on the C-BASE for this subtest, the MoGEA was designed to provide teacher candidates with prompts based in education to better assist them in formulating responses. Student responses would still be evaluated on form rather than content, however. The item validation and standard setting process for the other four subtests also occurred during the summer of 2013. Pearson sent an open invitation to K-12 educators and EPP personnel statewide to participate in this process in an attempt to gather a group of education professionals that would represent the entire state and reflect the state’s diversity. More than 1500 people took part in
reviewing the MoGEA and other exams being created by Pearson for use by the state. Once gathered, these professionals reviewed test items to determine if the “just acceptable candidate” should be able to answer the questions after completing the 42-hour general education course requirement. To ensure the validity and reliability of the MoGEA, the Revised Angoff Method was used to determine which test items should remain on the subtests and which should be removed. Through this process, the subtest frameworks were repeatedly reviewed and revised until agreed upon (G. Hairston, personal communication, September 18, 2014).

**Basic Knowledge Testing and Accountability**

Since the beginning of the formalization of free public education, candidates were expected to have an appropriate general knowledge base in order to qualify for the teaching profession. “In 1834, Pennsylvania was the first state to require teachers to pass exams on basic skill competencies of reading, writing and arithmetic” (Blackford et al., 2012, p. 6), and other states gradually followed suit. While it did not become common practice for states to formally assess teacher candidates’ basic knowledge until the 1960s (Goldhaber, 2007), a direct relationship between basic knowledge and teacher quality has long been assumed. During the 1990s, litigation in the state of California regarding content area exams required for teacher certification renewed the government’s focus on basic skill requirements for teachers. The use of basic knowledge tests to screen teachers became even more prevalent (Memory, Coleman, & Watkins, 2003), a trend that continued into the 21st century (Goodman, Arbona, & de Rameriz, 2008). As of 2012, in accordance with federal requirements for teacher certification, 42 states require passing scores on a basic knowledge test in order for teacher candidates to earn certification.
The historical record thus shows basic knowledge testing of teacher candidates has always been used as a means of accountability for teacher quality; whether they should be used for this purpose is the topic of much debate.

Goldhaber (2007) proposes the following reason states require minimum passing scores on licensure tests: “States implement teacher testing requirements in order to exclude individuals who would have been teachers of unacceptably low quality” (p. 768). He further posits:

The primary goal of a state’s licensure policy is to limit the number of individuals who would make for ineffective (and potentially harmful) teachers in public schools and a secondary goal is to create a teacher workforce of higher overall quality than would exist in its absence. (p. 767)

Goldhaber’s comments on the purposes of teacher certification and qualifying examinations illuminate the underlying assumptions regarding the use of basic knowledge testing to screen teacher candidates. In their research on this issue, Memory, Antes, Corey, and Chaney (2001) identify the following assumptions as “among the views motivating some individuals and organizations calling for higher qualifying scores on basic skills tests” (p. 182):

1. “Individuals with greater competency in reading, writing, and basic mathematics are more effective teachers than individuals with less competency in these areas” (p. 181-2).
2. “Certain minimal levels of competency in the basic skills are necessary to be an effective teacher” (p. 182).
Despite these assumptions regarding the relationship of the quality of teachers to the strength of their general base of knowledge, education professionals do not agree on whether basic knowledge tests are effective screening tools for teacher candidates or if they should serve as a means for holding teachers and EPPs accountable for teacher quality.

Proponents of the use of basic knowledge tests as a prerequisite for teacher licensure assert the content assessed on these exams is a reflection of the curricular content found in the nation’s public school systems (Gitomer, Brown, & Bonett, 2011). They believe classroom teachers should have a general base of knowledge that equips them to teach the same material to their students. Additionally, basic knowledge tests are often required to gain admittance into EPPs because they “are designed to represent skills that students in college need to have in order to navigate their coursework” (Gitomer et al., 2011, p. 441). Gitomer, Brown, and Bonett (2011) offer further support for the use of basic knowledge tests to screen teacher candidates:

The evidence is extremely consistent that the basic skills tests are more than an unnecessary obstacle for otherwise qualified and committed individuals.

Individuals who pass basic skills tests at borderline levels are far less likely to pass licensure tests than are candidates who meet the median state-level basic skills test requirements. Struggling with basic skills tests is a harbinger of later difficulty. (p. 441)

Whether used for measuring a teacher candidate’s general knowledge base or as a predictor of future success in passing state licensure examinations, basic knowledge tests are perceived by some to be legitimate assessment tools and accountability measures.
In contrast to those who advocate the testing of teacher candidates’ basic knowledge, many education professionals also negate their value in the process of earning teacher licensure. In their research of the aforementioned assumptions regarding the use of basic knowledge tests, Memory et al. (2001) reported their findings “provide no support for the…assumption that certain minimal levels of competency in the basic skills are necessary to be an effective teacher” (p. 188). Further, they explained their research “provides no firm support for the assumption that individuals with greater competency in reading, writing, and basic mathematics are more effective as teachers than individuals with less competency in these areas” (p. 182-3).

Additionally, in opposition to the findings of Gitomer et al. (2011), Henry et al. (2013) reported that teacher testing “may not strongly predict whether a student is successful in a [teacher preparation program]” (p. 443). The following statement made by the American Psychological Association also serves as a warning against the use of basic knowledge tests as a prerequisite for acceptance into EPPs: “Avoid using a single test score as the sole determinant of decisions about test takers. Interpret test scores in conjunction with other information about individuals” (American, 2004). Henry et al. (2012) argue similarly for the use a wider variety of accountability measures for EPPs than simply test scores. The use of basic knowledge testing as a means of holding individuals accountable for student success has gradually shifted beyond this intention; basic knowledge testing is now being used to hold institutions, specifically EPPs, accountable for the success of K-12 students (Feuerstein, 2011). Regardless of whom is being held accountable for K-12 student growth and performance, the related literature
and research indicate it is not advisable for basic knowledge tests, or any standardized tests, to be used for this purpose.

**Test Bias**

Much research has been conducted regarding the potential bias of basic knowledge tests and their use in the screening process for teacher candidates. Wakefield (2008) warned that “high-stakes tests...guard the door to the teaching profession in many of the nation’s states and territories” (p. 380). Gitomer et al. (2011) took a similar stance and furthered the argument against the use of basic knowledge tests as qualifying examinations for teacher certification by discussing their effect on minority candidates:

Significant numbers of individuals taking basic skills tests do not pass, making them ineligible to continue in teacher education. A disproportionately large number of African American candidates do not pass the tests, leading some to conclude that basic skills tests unfairly restrict the opportunities of minority students who want to enter the teaching profession. Being unfair, critics argue that without such basic skills testing requirements, these same teacher candidates would be able to successfully complete their teacher preparation and go on to be successful teachers. (p. 431)

In their study involving basic knowledge tests used as a qualifier for admission into EPPs, Gitomer, Latham, and Ziomek (1999) found an alarming disparity in the passing rates of students from different demographics: “Passing rates for African American candidates were approximately 52%, whereas White candidates passed at a rate of approximately 87%” (Gitomer et al., 2011, p. 432). The door to the teaching profession
of which Wakefield (2008) wrote is therefore closed to African American teacher candidates much sooner and much more frequently than it is to their white counterparts.

It is not only the scores minority students receive on basic knowledge tests that serve to eliminate them from teacher candidacy. Wakefield (2006) explained that the fees required for simply taking these tests, numerous times in many cases, deter students from entering EPPs: “low-income candidates, who may coincidentally be minorities or recent immigrants, are hit hardest by screening tests. Cost alone may force many to opt out of the teaching profession” (p. 82). Wakefield (2006) elaborated:

In addition, candidate screening tests are costly. Many students and their parents save for years, work extra jobs, and cut personal expenses in an effort to earn their college degrees. A significant portion of the population will pay their fees and pass teacher-screening tests without any added effort. For those who can’t, the additional testing, preparation, and travel expenses can approach the cost of a semester’s tuition at a state school. (p. 82)

Whether it be from the cost of these tests or the scores received on them, basic knowledge examinations are clearly eliminating many minority students from the teaching profession, particularly in the African American and Hispanic populations (Angrist & Guryan, 2004).

When discussing the issue of test bias and its relationship to minority teacher candidates, it is critical to note why the elimination of these candidates from the education profession is problematic. Of primary importance is the fact that statistically the demographics of the nation’s student population are changing dramatically while the teacher population is not. Selwyn (2007) explained that in 1999-2000, “74% of public
school teachers [were] women, and 84% of the teacher corps White” (p. 125). He noted, in contrast to teacher demographics, “the numbers show that percentages of students of color and students whose first language is not English are rising….by 2050 students of color will constitute 57% of the total student population in public schools” (p. 126).

Many assume that highly qualified teachers, regardless of their demographics, will be effective teachers for any student population. Research has indicated otherwise, however. Gitomer et al. (2011) summarized the issue:

> There are multiple reasons to be concerned about these patterns of demographic representation of teachers. The fundamental argument is that there is social and educational utility to be gained by having a greater presence of teachers who share common personal histories with their students and who can more closely identify with societal and persona challenges faced by students. At the same time, students are more likely to identify and respond to such teachers. (p. 431)

Gitomer et al. (2011) also pointed out that, particularly for the African American population, students have better attitudes toward school, achieve more in the classroom, have lower absentee rates, and are perceived to work harder when they are instructed by a teacher with the same demographic background. Unfortunately, Wakefield (2008) warned, “under the high-stakes screening tests, minorities can expect to see fewer teachers of their own race teaching their children” (p. 386).

Despite the well-documented benefits of having minority teachers in classrooms and the need for an increase in the number of minority teachers throughout the country, research repeatedly shows that the testing of teacher candidates’ basic knowledge serves to decrease the diversity of the teaching workforce (Goodman et al., 2008). Selwyn
(2007) explained, “the more pressure there is for teacher applicants to take and pass paper-and-pencil teacher tests, the more it works against diversifying the teacher corps” (p. 127). This is not surprising, for, as Wakefield (2006) reported, “demographic studies produced by standardized testing companies suggest that well over half the minority/low-income populations would be rejected by screening tests” (p. 82). Many would-be minority teacher candidates might choose not even to consider the field of teaching because of the current high-stakes testing requirements. Memory et al. (2001) contended that simply hearing of the costs and difficulties associated with taking and passing basic knowledge tests deters minority candidates from pursuing a career in education.

Although much research conducted by educational professionals provides clear evidence that basic knowledge tests used to screen teacher candidates are biased against minority populations, especially Hispanic and African American students, past and current trends in education reform indicate their use will continue and possibly even increase. Especially due to the ever-present demand to supply public schools with higher quality teachers, many states have responded to this call by increasing not only the use of basic knowledge tests but the cut scores students must achieve as well (Memory et al., 2003). Furthering the bias minority students already experience when taking such tests, raising cut scores has a markedly more significant effect on these students than it does on their white counterparts. Memory et al. (2003) reported the results of a study that examined the effect a one-point increase in the cut scores on each of the reading, writing, and mathematics subtests of a basic knowledge test would have for African American students’ elimination from EPPs. The results were alarming:
• A one-point increase in the cut score for the reading subtest would result in the elimination of an additional 5% of the African American teacher candidates.

• A one-point increase in the cut score for the writing subtest would result in the elimination of an additional 9% of the African American teacher candidates.

• A one-point increase in the cut score for the mathematics subtest would result in the elimination of an additional 4% of the African American teacher candidates.

What is particularly troubling about these findings is that the students who would take this exam have already demonstrated academic success as they are usually one or two years into their collegiate studies and have, in some cases, already earned an undergraduate degree. Additionally, they would also have already expressed the desire to become educators by declaring education as their major and through their willingness to pay the fee to take the basic knowledge test (Memory et al., 2003).

The issue of test bias against minority teacher candidates is of obvious concern to the teaching profession. Whether because basic knowledge tests serve to deter would-be teachers from even considering the profession or because they eliminate teacher candidates already in pursuit of teaching certificates, it is clear basic knowledge tests play a role in determining the demographic makeup of the country’s teaching corps. Education professionals are not in agreement regarding whether the achievement of state-mandated qualifying scores on basic knowledge tests ensures the quality or effectiveness of teachers in the classroom; however, there is much evidence indicating the use of basic knowledge tests as a requirement for certification negatively affects the diversification of the teacher
workforce. Despite this evidence, it appears the use of tests of basic knowledge to screen teacher candidates will continue, at least in the foreseeable future.

**Systemic Change**

When examined through the lens of this study’s conceptual framework, the three topics previously discussed in this chapter represent numerous attempts at systemic change. Historical events such as the Great Depression, World War II, and the Soviet Union’s launch of Sputnik served as catalysts for changes in teacher certification requirements as teacher supply and demand fluctuated or the nation sought to improve the quality of education its student received. The acts of the federal government previously described also intentionally or indirectly affected teacher training and licensure. As is one of the tenets of systemic change, attempts to change one aspect of the system—NCLB’s requirement that teachers be highly qualified, for example—actually affect all the system’s components throughout the change process. Government leaders might desire stronger teachers in order to improve student achievement, but the effects of legislation created to improve the quality of education are far-reaching. When state boards of education and other government agencies receive federal mandates regarding teacher certification, stakeholders at all levels are affected: colleges and universities; the faculty, students, and administration of EPPs; public school administrators and teachers; and the families and students served by public schools.

As previously mentioned, federal legislation has also directly influenced the public perception of teachers and education in the United States, a byproduct of the systemic change process that might not necessarily be as quantifiable as test scores but is
certainly of chief importance to the education profession. Peck, Gallucci, and Sloan (2010) explained:

Changes in federal policy within the United States affecting teacher education have emerged amid a rhetoric of “crisis”, “risk”, and “failure,” as Congress has enacted intensified reporting and accountability mandates aimed at increased control of teacher education policy and practice….A major challenge for teacher educators is how to negotiate programmatic responses to new state and federal mandates in the context of this negative rhetorical climate… (p. 452)

Further, legislative attempts at change might even be counterproductive due to the public’s perception:

The meanings teachers construct around reform policy often affect their commitment to their work. The underlying dilemma here is that external policy mandates, particularly when accompanied by negative rhetoric, may undermine the very motivational qualities necessary to their successful implementation. (p. 452)

Attempts at educational reform at the national or state level undoubtedly affect stakeholders throughout the educational system. However, legislative attempts at educational change have historically left out many of the stakeholders that most immediately felt the impact of federal mandates when decisions are being made. Joseph (2004) warned against this: “Broad stakeholder ownership is the fundamental bedrock upon which all other aspects of the conceptual framework of the systemic change process are built” (p. 7). Commitment to participation, one of six identified process values of the systemic change process, calls for the involvement of all stakeholders when making
decisions regarding systemic change (Joseph et al., 2002). Unfortunately, federally mandated requirements regarding teacher training and certification have repeatedly been written and passed without the involvement of those who are not only most directly affected by their stipulations but who are also responsible for their implementation.

The use of basic knowledge tests and their bias against minority students also represents an issue regarding the process of systemic change particularly when minority teacher candidates, public school students, and parents are considered. These groups are all stakeholders in the nation’s educational system, yet their needs oftentimes appear to go unnoticed when changes to the process of teacher certification are created, considered, and implemented. In fact, as previously explained in this chapter, federal- and state-mandated testing of teacher candidates is actually counterproductive to meeting the needs of minority populations. Not only are these stakeholders not being involved in the decision making process, but research shows the decisions being made negatively affect the educational experiences of this population.

Finally, the decision of primary importance to this study—the decision to change from the C-BASE to the MoGEA as the qualifying basic knowledge test for admission to EPPs in the state of Missouri—can be analyzed through the lens of systemic change. As previously discussed in the first chapter of this study, Banathy identified six process values that provide the foundation for systemic change: commitment to participation, commitment to idealized design, design is continuous, nurturing human values and human quality, design community, and design conversation. Each of these values can be used to further explore the decision made by Missouri’s BOE.
Commitment to participation, the first process value, addresses the importance of involving all stakeholders in every aspect of the change process and advises that those responsible for implementing the changes also be the same parties responsible for designing the process (Joseph et al., 2002). Once again, the decision to change to the MoGEA was made without the involvement of EPP personnel. It is the educators and administrators within Missouri’s EPPs who are responsible for preparing teacher candidates for the MoGEA exam through direct instruction and curriculum design. While K-12 educators and EPP personnel were involved in the creation of the MoGEA, the original decision was made without their input. This decision was made, however, with the intention of improving Missouri’s process of data collection and reporting as mandated by RttT, demonstrating an awareness of the second process value of systemic change, commitment to idealized design.

The third process value, design should be continuous, is represented by the ongoing work being done by Dr. Hairston’s office in Missouri with Pearson’s Evaluation Systems Group. During the 2013-2014 academic year, Dr. Hairston met regularly with this group to discuss feedback he received from several stakeholders regarding all aspects of the MoGEA. These regular meetings demonstrate a desire to continuously improve the implementation of the BOE’s decision (G. Hairston, personal communication, September 18, 2014). In contrast, nurturing human values and human quality, the fourth process value, was inconsistently observed throughout the systemic change process of concern to this study. While the BOE’s involvement of various stakeholders in the MoGEA design and revision process and the meetings Dr. Hairston held regularly with Pearson’s Evaluation Systems Group represent a commitment to supporting those involved in the
implementation of the decision, these same groups were not consulted when the original
decision was made.

Similar to the manner in which the fourth process value was observed during the
decision to change from C-BASE to the MoGEA, Banathy’s fifth and sixth process
values, design community and design conversation, were only represented by the actions
of Missouri’s BOE after this decision had already been made. EPP personnel, while not
involved in the original decision, are the “users of the system being created” (Joseph et
al., 2002, p. 385) and were only included in the design and review portion of the process.

In summary, this important decision affects the entire system of teacher education
in the state of Missouri and was made by only a few key stakeholders. Joseph (2004)
advised: “Bringing stakeholders together…strengthens the change process. If voices of
stakeholders are left out…then the change process is weakened and is more susceptible to
adverse reactions from these very same stakeholders” (p. 8). While EPP faculty and
administrators, as well as teachers throughout the state of Missouri, were involved in the
item validation and standard setting process for the MoGEA exam, these important
parties were not included in the original decision to change exams (G.H. Hairston,
personal communication, September 18, 2014). As these same stakeholders—EPP
students, faculty, and administrators—are now preparing for the MoGEA and the
potential effect it will have on student performance and admission to EPPs, time will tell
if this particular attempt at systemic change will prove successful despite Joseph’s
forewarning.
CHAPTER THREE
RESEARCH DESIGN AND METHODOLOGY

The decision to use the Missouri General Education Assessment (MoGEA) rather than the College Basic Academic Subjects Examination (C-BASE) is a policy change that affects all stakeholders throughout Missouri’s educational system. Because the MoGEA was a new examination and there were no prior student results on which to base cut scores for the 2013-2014 school year, educator preparation programs (EPPs) were individually responsible for setting their own cut scores for the first two years of the exam’s implementation. The current availability of 2013-2014 MoGEA results afforded the researcher the opportunity to satisfy the desire of the Missouri Department of Elementary and Secondary Education (DESE) and EPP administrators to recommend statewide subtest cut scores and determine whether any student demographic characteristics can be used to predict success on the MoGEA.

This chapter will explain the research design and methodology for the study. After first reviewing the problem, purpose, research questions, and null hypotheses associated with the research, the methodology will then be described. The design of the research will next be explained, including the independent and dependent variables for the study. A description of the process used to collect data for the study will follow, and the steps taken to ensure the confidentiality of the study group members and their respective institutions will be highlighted. The instrument of use in the research is then explained prior to an in-depth discussion of the data analysis procedures associated with each of the study’s research questions presented in both narrative form and in Table 2.
**Problem Statement**

At the time of this study’s proposal, no statewide qualifying scores existed for the Missouri General Education Assessment. In order for DESE to establish examination cut scores based on the input of Missouri’s EPP personnel and inform the final decision made by the State Board of Education (BOE), it was critical that statewide results of the 2013-2014 MoGEA be collected and analyzed to ensure a more representative population of test scores be used to establish valid and reliable cut scores for use in 2015.

**Research Purpose**

The purpose of this quantitative study was to recommend statewide cut score norms for the MoGEA to the executive board of the Missouri Association of Colleges for Teacher Education (MACTE) for the purpose of informing its representatives to the Missouri Advisory Council of Certification for Educators (MACCE). Additionally, this study aimed to identify demographic characteristics of EPP candidates that might serve as predictors of success on the MoGEA. If predictor variables are determined, EPP administrators can identify which of their education students will most likely require additional support and preparation for the MoGEA. This information will also make DESE and the BOE aware of which EPP candidate demographics are likely to be denied EPP admission.

**Research Questions and Null Hypotheses**

This study was designed to address the following research questions and their accompanying null hypotheses:

1. What are the descriptive summary statistics for each MoGEA subtest for the total population of the study group and when broken out by institutional selectivity?
2. Do statewide scores on the MoGEA subtests for the 2013-2014 academic year represent a normal distribution?

H₀₁: Statewide scores on the Missouri General Education Assessment for the 2013-2014 academic year do not represent a normal distribution.

3. What are the mean and median scores for each decile within each MoGEA subtest?

4. When composite ACT score is used as a covariate, is there a difference in MoGEA mean subtest scores for the following demographic variables?
   a. Ethnic origin
   b. Gender
   c. Total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained
   d. Cumulative grade point average at the time the student’s highest reported score on the MoGEA was obtained
   e. Classification of the candidate as receiving financial aid or not receiving financial aid

H₀₂: There is no difference in MoGEA mean subtest scores between demographic variables.

5. How will establishing cut scores at the lowest score within each decile affect teacher candidate admission for populations that have historically been shown to be negatively affected by test bias?
H₀₃: Establishing cut scores at the lowest score within each decile has no effect on teacher candidate admission for populations that have historically been shown to be negatively affected by test bias.

6. Is it possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study?

H₀₄: It is not possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study.

**Methodology**

This section will describe the research design, population and sample size, data collection process, instrumentation, and data analysis for this quantitative study. The chosen methodology was selected to obtain the data necessary to answer the study’s research questions.

**Research Design**

Based on the research questions of this project and the methods required to answer them, a quantitative non-experimental approach was designed. Additionally, because this study aimed to “identify and assess the causes that influence outcomes”, specifically in regard to the predictability of students achieving qualifying scores on the Missouri General Education Assessment, a postpositivist philosophical worldview was supported (Creswell, 2009, p. 7). The postpositivist worldview best suited this project because the study’s findings were based on “data, evidence, and rational considerations”, and it sought to “explain the situation of concern” and “describe the causal relationships of interest” (Creswell, 2009, p. 7).
Independent Variables

For this study, the following independent variables were examined for each educational preparation program (EPP) candidate:

- Selectivity of institution. The four possible classifications for selectivity were open enrollment, moderately selective, selective, and highly selective.
- Composite ACT score.
- Ethnicity.
- Gender.
- Total number of earned credit hours at the time the student’s highest reported score on the MoGEA was obtained.
- Cumulative college-level grade point average at the time the student’s highest reported score on the MoGEA was administered.
- Classification of the candidate as receiving financial aid or not receiving financial aid.
- Class level. The possible classifications were freshman, sophomore, junior, senior, and graduate student.

Dependent Variables

The dependent variables examined in this study were the highest scores obtained by educator preparation program (EPP) candidates on each of the five MoGEA subtests: English language arts, writing, math, science, and social studies.

Study Group

The subjects included in the study group for this project were EPP candidates throughout the state of Missouri who had taken the MoGEA between September 1, 2013,
and September 1, 2014. MoGEA subtest scores were obtained from 10 EPPs throughout Missouri creating a sample size of 1037 EPP candidates. As of September 1, 2014, 4996 students had taken the MoGEA (G. Hairston, personal communication, September 18, 2014). With a total population of 4996 at a 95% confidence level and a confidence interval of 5%, a sample size of 357 would be necessary to accurately represent the population (Sample Size Calculator, n.d.). Therefore, the sample size of 1037 scores collected for this study can be considered an accurate representation of the population (Field, 2009).

**Data Collection**

Archival MoGEA test score data and specific independent variables from EPP candidates were collected to answer the research questions of this study. In order to obtain EPP candidate data, an email containing an Excel spreadsheet was sent to a list of self-selected EPP administrators who expressed interest in participating in the study. EPP personnel at these colleges and universities filled in the spreadsheet with their respective institutional data, stripping any personally identifiable information from individual student MoGEA subtest scores. Non-respondents were sent follow-up emails and, in some cases, personal telephone calls. Ultimately, data were collected from 10 institutions resulting in the aforementioned sample size of 1037 students. Upon receipt of each participating EPP’s data, identifying information for the institution was removed from the data by a third party and nominal or interval values were assigned to the various independent variables of importance to the study to protect each institution’s anonymity (Field, 2009).
As is noted in the limitations addressed in Chapter 1 of this study, the availability of student information regarding the demographic variables of concern to the study’s sixth research question varied from institution to institution. The minimum requirement for inclusion of a subject in this study was a score on each of the MoGEA subtests. Beyond this baseline, the sample size for statistical analyses for each of the demographic variables differed. The specific sample size for each statistical analyses is addressed in Chapter 4.

Confidentiality and Data Anonymity Assurances

Throughout the data collection process, the confidentiality of EPP candidates was ensured. All personally identifiable information was stripped from individuals’ MoGEA subtest scores prior to their receipt by the aforementioned third party. For this reason, this study did not qualify as human subjects research and, subsequently, did not require International Review Board approval from the University of Missouri-Columbia (see Appendix A).

In addition to the protection of the identities of individual EPP candidates, the researcher also ensured the confidentiality of each EPP. Once an Excel document was completed by an EPP administrator and returned to the third party, the institution’s identifying information was stripped from the document and replaced with a nominal value identifying the institution by level of selectivity. This process ensured that no MoGEA subtest scores could be directly linked to the specific institution from which they originated.
Instrumentation

The instrument for this study was the Missouri General Education Assessment. Because no state-level, aggregated or disaggregated data analysis had been conducted on the results of the MoGEA prior to this study, the researcher designed research questions to analyze descriptive statistics for the study group’s MoGEA subtest scores and to determine whether different demographic populations scored differently on the subtests.

Data Analysis

The purpose of the selected data analyses was to answer the research questions previously discussed in this chapter. Statistical analyses were conducted using both SPSS® statistical analysis software and Microsoft Excel® spreadsheet software. As they pertained to the specified research questions, the following statistical analyses were conducted:

• Research Question 1: In order to generate descriptive summary statistics, the researcher found the mean, median, mode, and standard deviation for each MoGEA subtest. The researcher performed these tests for the total population of the study group and after breaking out the study group by institutional selectivity.

• Research Question 2: In order to determine whether student scores on the MoGEA represent a normal distribution, the researcher performed the Kalmogorov-Smirnov test for each subtest. The researcher performed this test for the total population of the study group and after breaking out the study group by institutional selectivity. Additionally, the researcher examined both histograms and Q-Q plots for each subtest, for the total population of the study group and when broken out by institutional selectivity.
• Research Question 3: The researcher performed decile analysis in order to
determine the mean and median scores for each decile within each MoGEA for
the total population of the study group.

• Research Question 4: The researcher originally proposed the performance of an
analysis of covariance (ANCOVA) with composite ACT scored used as a
covariate in order to determine whether differences existed in the mean scores on
each subtest for the various demographic variables specified for this research.
When the results of a preliminary ANOVA showed MoGEA subtest scores were
significantly dependent on composite ACT scores, an ANCOVA was not
performed (Field, 2009). The researcher then conducted additional ANOVA tests
for teacher candidate ethnicity and class level and independent-samples \( t \)-tests for
ethnicity, gender, and financial aid.

• Research Question 5: The researcher performed descriptive analysis for each
MoGEA subtest in order to determine whether establishing cut scores at the
lowest score in each decile affects teacher candidate admission by demographic
characteristic.

• Research Question 6: The researcher first computed a Pearson Product-Moment
Correlation Coefficient (PPMCC) to determine whether the study’s dependent
variables, teacher candidates’ scores on the five MoGEA subtests, were
significantly correlated with the following demographic characteristics of the
study’s population: total number of credit hours earned at the time the student’s
highest reported score on the MoGEA was obtained, cumulative GPA at the time
the student’s higher reported score on the MoGEA was obtained, and maximum
composite ACT score obtained by the teacher candidate (Field, 2009). Following the PPMCC, multivariate analysis of variance (MANOVA) and discriminant function analysis were performed in order to determine whether it is possible for to predict whether a student will receive qualifying scores on the MoGEA based on the study’s identified independent variables.

Table 2 displays the statistical analyses strategies for the study by research question.

Table 2

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Analysis Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Descriptive summary statistics</td>
<td>Mean, median, mode, standard deviation</td>
</tr>
<tr>
<td>2. Distribution</td>
<td>Kalmogorov-Smirnov and data displays</td>
</tr>
<tr>
<td>3. Decile analysis</td>
<td>Mean and median</td>
</tr>
<tr>
<td>4. Demographic variable differences</td>
<td>ANCOVA, ANOVA*, independent-samples t-test*</td>
</tr>
<tr>
<td>5. Cut score effect on demographic variables</td>
<td>Descriptive analysis</td>
</tr>
<tr>
<td>6. Predictor variables</td>
<td>PPMCC, MANOVA, and discriminant functional analysis</td>
</tr>
</tbody>
</table>

* These statistical analyses were not originally proposed but were used when the results of the preliminary ANOVA determined an ANCOVA was unnecessary.

Summary

The desire for a greater understanding of student performance on the Missouri General Education Assessment during the 2013-2014 school year indicated a need for this study. To gain this comprehensive perspective required the collection and analysis of MoGEA subtest scores from EPPs throughout the state. The call for a recommendation of statewide cut scores fueled the study’s research questions, which necessitated the accompanying data analysis procedures previously described. Finally, so that EPP faculty and administrators might proactively identify teacher candidates prone to struggle with
earning qualifying scores on the MoGEA based on each student’s individual demographics, further data analyses were performed.
CHAPTER FOUR
PRESENTATION AND ANALYSIS OF DATA

This chapter provides the presentation and analysis of the data collected to test the research questions and hypotheses of the study. A brief review of the problem and purpose of the study are first provided, along with a review of the instrument of the study. Next, an overview of the organization of the data analysis is outlined, and the demographic characteristics of the study group are presented in table format. Research questions and hypotheses are then broken down in the analysis of data showing the results of each statistical test used in determining the acceptance or rejection of the null hypotheses, where applicable. Finally, a summary of the chapter is included.

Review of Problem and Purpose of the Study

This study was necessitated by the lack of information and analysis of statewide scores on the Missouri General Education Assessment (MoGEA) during its first year of use, the 2013-2014 academic year. As previously discussed, Missouri’s State Board of Education (BOE) requested feedback from teacher preparation program administrators and instructors regarding student performance on the MoGEA in order to set statewide cut scores beginning in the 2015-2016 academic year, and this study served to satisfy this request. The purpose of this quantitative study was to recommend statewide cut score norms for the MoGEA to the executive board of the Missouri Association of Colleges for Teacher Education (MACTE) for the purpose of informing its representatives to the Missouri Advisory Council of Certification for Educators (MACCE). Additionally, this study aimed to identify demographic characteristics of educator preparation program (EPP) candidates that might serve as predictors of success on the MoGEA. The
conceptual underpinning for the study, the process of systemic change, served to guide the researcher throughout the scope of the entire project and provided a lens through which the topic could be critically examined.

**Instrumentation Review**

The instrument for this study was the Missouri General Education Assessment. Created by Pearson’s Evaluation Systems Group to serve as the basic knowledge assessment for candidates for EPPs throughout the state of Missouri (G. Hairston, personal communication, September 18, 2014), the MoGEA consists of five subtests: English language arts, writing, math, science, and social studies. In order to gain acceptance into an EPP, students must obtain minimum cut scores on each of the subtests. Student subtest scores, as well as demographic information for individual students and participating institutions, were collected by EPP personnel, stripped of all personally-identifiable information, and submitted to a third party data collector. This party then stripped each institution’s identifying information prior to submitting the data to the researcher. Through this process, the confidentiality of teacher candidates whose scores were submitted for the study was ensured because individual scores could not be linked to specific students or institutions. Therefore, this study did not constitute human subjects research and was exempt from International Review Board approval (see Appendix A).

**Organization of Data Analysis**

The population of the study group and presentations of the demographic characteristics of the study group are provided in both narrative form and tables. In addition, data and the statistical analysis for each research question and its corresponding null hypothesis are examined. The statistical tests used included the following:
descriptive summary statistics for research question one (RQ1), a Kolmogorov-Smirnov (K-S) distribution test and analysis of data displays for research question two (RQ2), examination of the mean and median of each decile for research question three (RQ3), one-way analysis of covariance (ANCOVA), one-way analysis of variance (ANOVA), and independent-samples $t$-tests for research question four (RQ4), general descriptive analysis for research question five (RQ5), and Pearson Product-Moment Correlation Coefficient (PPMCC), multivariate analysis of variance (MANOVA), and discriminant function analysis for research question six (RQ6).

**Population of the Study Group**

The total population consisted of subtest scores collected from 1037 individual EPP candidates who took at least one of the MoGEA subtests between September 1, 2013, and September 1, 2014. With a total population of 4996 students taking the MoGEA during this time period, at a 95% confidence level and a confidence interval of 5%, a sample group of 357 EPP candidates would be required for the sample to be considered statistically representative of the population; the 1037 collected can therefore be considered an accurate representation of the study group’s total population (Field, 2009).

**Demographic Characteristics of the Study Group**

Subtest scores were collected from ten EPPs throughout the state of Missouri. In total, these ten EPPs submitted subtest scores for 1037 individual EPP candidates. The demographic characteristics presented in Tables 3-10 are those that were of importance to the study’s research questions and were considered independent variables for the study.
Each table displays the frequency of teacher candidates for each category and presents this number as a percent of the total 1037.

Table 3 displays the participating institutions broken down by institutional selectivity.

Table 3

*Institutional Selectivity of Participating EPPs*

<table>
<thead>
<tr>
<th>Institution Type</th>
<th>Frequency</th>
<th># Candidates</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Enrollment</td>
<td>1</td>
<td>12</td>
<td>1.2</td>
</tr>
<tr>
<td>Moderately Selective</td>
<td>5</td>
<td>485</td>
<td>46.8</td>
</tr>
<tr>
<td>Selective</td>
<td>4</td>
<td>540</td>
<td>52.1</td>
</tr>
</tbody>
</table>

Tables 4-10 display the demographic characteristics for the 1037 EPP candidates whose subtest scores were submitted for the study. The demographic characteristics presented are those that were of importance to the study’s research questions and were considered the independent variables for the study. Each table displays the frequency of respondents for each category and also expresses this frequency as a percent of the total 1037 candidates. Table 4 displays the self-reported ethnicities of the 1037 teacher candidates whose scores were submitted for the study.

Table 4

*Teacher Candidates – Ethnicity*

<table>
<thead>
<tr>
<th>Self-Reported Ethnicity</th>
<th>Frequency</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Indian/Alaskan Native</td>
<td>11</td>
<td>1.1</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>7</td>
<td>0.7</td>
</tr>
<tr>
<td>Black, Non-Hispanic</td>
<td>52</td>
<td>5.0</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>28</td>
<td>2.7</td>
</tr>
<tr>
<td>Non-Resident Alien</td>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>Bi- or Multi-Racial</td>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>White, Non-Hispanic</td>
<td>757</td>
<td>73.0</td>
</tr>
<tr>
<td>Unknown/Not Reported</td>
<td>178</td>
<td>17.2</td>
</tr>
</tbody>
</table>
Table 5 displays the gender of the teacher candidates whose scores were submitted for the study.

Table 5

<table>
<thead>
<tr>
<th>Gender</th>
<th>Frequency</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>790</td>
<td>76.2</td>
</tr>
<tr>
<td>Male</td>
<td>223</td>
<td>21.5</td>
</tr>
<tr>
<td>Not Reported</td>
<td>24</td>
<td>2.3</td>
</tr>
</tbody>
</table>

Table 6 displays the total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained for the teacher candidates whose scores were submitted for the study, and Figure 1 displays the same information as a histogram.

Table 6

<table>
<thead>
<tr>
<th>Total Credit Hours</th>
<th>Frequency</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 42</td>
<td>105</td>
<td>12.6</td>
</tr>
<tr>
<td>42 – 51</td>
<td>77</td>
<td>9.3</td>
</tr>
<tr>
<td>52 – 61</td>
<td>101</td>
<td>12.0</td>
</tr>
<tr>
<td>62 – 71</td>
<td>115</td>
<td>13.8</td>
</tr>
<tr>
<td>72 – 81</td>
<td>100</td>
<td>12.0</td>
</tr>
<tr>
<td>82 – 91</td>
<td>95</td>
<td>11.4</td>
</tr>
<tr>
<td>92 – 101</td>
<td>60</td>
<td>7.1</td>
</tr>
<tr>
<td>102 – 111</td>
<td>51</td>
<td>6.0</td>
</tr>
<tr>
<td>112 – 121</td>
<td>42</td>
<td>5.0</td>
</tr>
<tr>
<td>122 – 131</td>
<td>36</td>
<td>4.3</td>
</tr>
<tr>
<td>≥ 132</td>
<td>55</td>
<td>6.6</td>
</tr>
<tr>
<td>Not Reported</td>
<td>200</td>
<td>19.3</td>
</tr>
</tbody>
</table>
Figure 1. Total number of credit hours earned by teacher candidates.

Table 7 displays the cumulative GPAs at the time their highest MoGEA scores were obtained for the teacher candidates whose scores were submitted for the study, and Figure 2 displays the same data as a histogram.

Table 7

<table>
<thead>
<tr>
<th>GPA</th>
<th>Frequency</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 2.40</td>
<td>24</td>
<td>2.3</td>
</tr>
<tr>
<td>2.40 – 2.59</td>
<td>36</td>
<td>3.5</td>
</tr>
<tr>
<td>2.60 – 2.79</td>
<td>130</td>
<td>12.5</td>
</tr>
<tr>
<td>2.80 – 2.99</td>
<td>120</td>
<td>11.6</td>
</tr>
<tr>
<td>3.00 – 3.19</td>
<td>136</td>
<td>13.1</td>
</tr>
<tr>
<td>3.20 – 3.39</td>
<td>143</td>
<td>13.8</td>
</tr>
<tr>
<td>3.40 – 3.59</td>
<td>135</td>
<td>13.0</td>
</tr>
<tr>
<td>3.60 – 3.79</td>
<td>98</td>
<td>9.5</td>
</tr>
<tr>
<td>3.80 – 4.00</td>
<td>103</td>
<td>9.9</td>
</tr>
<tr>
<td>Not Reported</td>
<td>112</td>
<td>10.8</td>
</tr>
</tbody>
</table>
**Figure 2.** Cumulative GPAs of the total population of the study at the time their highest MoGEA scores were obtained.

Table 8 displays the total population of the study broken down by whether or not they received financial aid.

Table 8

<table>
<thead>
<tr>
<th>Teacher Candidates – Receive Financial Aid?</th>
<th>Frequency</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>175</td>
<td>16.9</td>
</tr>
<tr>
<td>Yes</td>
<td>317</td>
<td>30.6</td>
</tr>
<tr>
<td>Not Reported</td>
<td>545</td>
<td>52.6</td>
</tr>
</tbody>
</table>
Table 9 displays the total population of the study broken down by class level.

**Table 9**

*Teacher Candidates – Class Level*

<table>
<thead>
<tr>
<th>Class Level</th>
<th>Frequency</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshman</td>
<td>181</td>
<td>17.5</td>
</tr>
<tr>
<td>Sophomore</td>
<td>238</td>
<td>23.0</td>
</tr>
<tr>
<td>Junior</td>
<td>271</td>
<td>26.1</td>
</tr>
<tr>
<td>Senior</td>
<td>170</td>
<td>16.4</td>
</tr>
<tr>
<td>Graduate Student</td>
<td>4</td>
<td>0.4</td>
</tr>
<tr>
<td>Not Reported</td>
<td>173</td>
<td>16.7</td>
</tr>
</tbody>
</table>

Table 10 displays the total population of the study broken down by their maximum ACT composite scores, and Figure 3 displays the same information as a histogram.

**Table 10**

*Teacher Candidates – Maximum ACT Composite Score*

<table>
<thead>
<tr>
<th>ACT Score</th>
<th>Frequency</th>
<th>Percent of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 20</td>
<td>146</td>
<td>14.1</td>
</tr>
<tr>
<td>20-21</td>
<td>153</td>
<td>14.8</td>
</tr>
<tr>
<td>22-23</td>
<td>149</td>
<td>14.4</td>
</tr>
<tr>
<td>24-25</td>
<td>95</td>
<td>9.2</td>
</tr>
<tr>
<td>26-27</td>
<td>82</td>
<td>7.9</td>
</tr>
<tr>
<td>28-29</td>
<td>47</td>
<td>4.5</td>
</tr>
<tr>
<td>≥ 30</td>
<td>26</td>
<td>2.5</td>
</tr>
<tr>
<td>Not Reported</td>
<td>339</td>
<td>32.7</td>
</tr>
</tbody>
</table>
Research Questions and Null Hypotheses

In response to the data collected, and after consultation with the researcher’s dissertation advisors, the researcher amended some research questions. As discussed in previous chapters of the study, data collection was timely; it was therefore necessary to terminate data collection at a time that still permitted the researcher to achieve the original purposes of the study. Despite repeated attempts to solicit participation in the study, some EPPs in Missouri did not submit teacher candidate MoGEA scores, and analyses regarding certain demographic characteristics and university classifications was not possible. Further issues regarding data collection, changes made to the original research questions, and recommendations for further data collection and research are addressed in Chapter Five.
Framed by the perspectives of the problem and purpose of the study, listed below are the research questions and hypotheses that guided the study.

1. What are the descriptive summary statistics for each MoGEA subtest for all study group members broken out by institutional selectivity?

2. Do statewide scores on the MoGEA subtests for the 2013-2014 academic year represent a normal distribution?
   
   $H_0_1$: Statewide scores on the MoGEA for the 2013-2014 academic year do not represent a normal distribution.

3. What are the mean and median scores for each decile within each MoGEA subtest?

4. When composite ACT score is used as a covariate, is there a difference in MoGEA mean subtest scores for the following demographic variables?
   
   a. Ethnic origin
   
   b. Gender
   
   c. Total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained
   
   d. Cumulative grade point average at the time the student’s highest reported score on the MoGEA was obtained
   
   e. Classification of the candidate as receiving financial aid or not receiving financial aid
   
   f. Class level

$H_0_2$: There is no difference in MoGEA mean subtest scores between demographic variables.
5. How will establishing cut scores at the lowest score within each decile affect teacher candidate admission for populations that have historically been shown to be negatively affected by test bias?

$H_03$: Establishing cut scores at the lowest score within each decile has no effect on teacher candidate admission for populations that have historically been shown to be negatively affected by test bias.

6. Is it possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study?

$H_04$: It is not possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study.

**Analysis of Data**

A statistical analysis and discussion for each research question used for the study is provided; where applicable, the hypothesis determination for each question is also provided.

**Research Question 1**

1. What are the descriptive summary statistics for each MoGEA subtest for the total population of the study group and when broken out by institutional selectivity?

The summary statistics—mean, median, mode, and standard deviation—on each subtest for the total population of the study group are displayed in Table 11.
Table 11

**Summary Statistics by MoGEA Subtest – Total Population of the Study**

<table>
<thead>
<tr>
<th>Subtest</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>English &amp; LA</td>
<td>989</td>
<td>232</td>
<td>231</td>
<td>231</td>
<td>23.8</td>
</tr>
<tr>
<td>Math</td>
<td>1006</td>
<td>204</td>
<td>202</td>
<td>177</td>
<td>33.0</td>
</tr>
<tr>
<td>Science</td>
<td>990</td>
<td>223</td>
<td>220</td>
<td>232</td>
<td>30.6</td>
</tr>
<tr>
<td>Social Studies</td>
<td>987</td>
<td>219</td>
<td>220</td>
<td>220</td>
<td>31.1</td>
</tr>
<tr>
<td>Writing</td>
<td>989</td>
<td>226</td>
<td>220</td>
<td>247</td>
<td>30.6</td>
</tr>
</tbody>
</table>

Figures 4-8 display the data set for each subtest for the total population of the study group as histograms.

*Figure 4.* MoGEA English and Language Arts subtest scores for the total population of the study.
Figure 5. MoGEA Math subtest scores for the total population of the study.

Figure 6. MoGEA Science subtest scores for the total population of the study.
Figure 7. MoGEA Social Studies subtest scores for the total population of the study.

Figure 8. MoGEA Writing subtest scores for the total population of the study.
The summary statistics—mean, median, mode, and standard deviation—for each of the five subtests broken out by institutional selectivity are displayed in Table 12.

Table 12

<table>
<thead>
<tr>
<th>Subtest Summary Statistics by Institutional Selectivity</th>
<th>Moderately Selective</th>
<th>Selective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean, Median, Mode, SD</td>
<td>Mean, Median, Mode, SD</td>
</tr>
<tr>
<td>English &amp; LA</td>
<td>230, 231, 249, 25.2</td>
<td>234, 237, 231, 22.3</td>
</tr>
<tr>
<td></td>
<td>(n = 460)</td>
<td>(n = 522)</td>
</tr>
<tr>
<td>Math</td>
<td>203, 195, 177, 33.9</td>
<td>205, 202, 202, 31.9</td>
</tr>
<tr>
<td></td>
<td>(n = 455)</td>
<td>(n = 527)</td>
</tr>
<tr>
<td>Science</td>
<td>219, 220, 226, 31.5</td>
<td>227, 226, 202, 28.5</td>
</tr>
<tr>
<td></td>
<td>(n = 451)</td>
<td>(n = 520)</td>
</tr>
<tr>
<td>Social Studies</td>
<td>215, 214, 195, 32.0</td>
<td>222, 220, 220, 30.0</td>
</tr>
<tr>
<td></td>
<td>(n = 467)</td>
<td>(n = 519)</td>
</tr>
<tr>
<td>Writing</td>
<td>223, 220, 193, 30.4</td>
<td>229, 247, 247, 30.8</td>
</tr>
<tr>
<td></td>
<td>(n = 468)</td>
<td>(n = 519)</td>
</tr>
</tbody>
</table>

Despite repeated attempts to solicit participation in the study, data were collected from only one open enrollment institution and included subtest scores from only 13 teacher candidates; this sample size was not large enough to be considered statistically significant in comparison to the sample sizes for both moderately selective and selective institutions and was therefore not included in the analysis. No data were collected from highly selective institutions despite multiple attempts to solicit participation in the study. Figures 9-13 display the data set for each subtest for institutions with moderately selective enrollment as histograms.
Figure 9. MoGEA English and Language Arts subtest scores for teacher candidates from moderately selective institutions.

Figure 10. MoGEA Math subtest scores for teacher candidates from moderately selective institutions.
Figure 11. MoGEA Science subtest scores for teacher candidates from moderately selective institutions.

Figure 12. MoGEA Social Studies subtest scores for teacher candidates from moderately selective institutions.
Figure 13. MoGEA Writing subtest scores for teacher candidates from moderately selective institutions.
Figures 14-18 display the data set for each subtest for institutions with selective enrollment as histograms.

*Figure 14.* MoGEA English and Language Arts subtest scores for teacher candidates from selective institutions.
Figure 15. MoGEA Math subtest scores for teacher candidates from selective institutions.

Figure 16. MoGEA Science subtest scores for teacher candidates from selective institutions.
Figure 17. MoGEA Social Studies subtest scores for teacher candidates from selective institutions.

MoGEA Social Studies Subtest Scores - Selective Institutions

Mean = 222.09
Std. Dev. = 30.047
N = 516
Tables 11 and 12 and Figures 4-22 display the descriptive summary statistics—mean, median, mode, and standard deviation—on each subtest for the total population of the study group and when the study group is broken out by institutional selectivity. Specifically when the subtest scores are broken out by institutional selectivity, differences in the descriptive summary statistics between moderately selective and selective institutions are revealed. For all five subtests, the mean and median scores are higher for teacher candidates from selective institutions:

- On the English and Language Arts subtest, the mean and median scores for teacher candidates from moderately selective institutions are 230 and 231, respectively. The mean and median scores for teacher candidates from selective institutions are 234 and 237, respectively.
• On the Math subtest, the mean and median scores for teacher candidates from moderately selective institutions are 203 and 195, respectively. The mean and median scores for teacher candidates from selective institutions are 205 and 202, respectively.

• On the Science subtest, the mean and median scores for teacher candidates from moderately selective institutions are 219 and 220, respectively. The mean and median scores for teacher candidates from selective institutions are 227 and 226, respectively.

• On the Social Studies subtest, the mean and median scores for teacher candidates from moderately selective institutions are 215 and 214, respectively. The mean and median scores for teacher candidates from selective institutions are 222 and 220, respectively.

• On the Writing subtest, the mean and median scores for teacher candidates from moderately selective institutions are 223 and 220, respectively. The mean and median scores for teacher candidates from selective institutions are 229 and 247, respectively.

For all but the Writing subtest, the standard deviation of the scores is higher for teacher candidates from moderately selective institutions.

• On the English and Language Arts subtest, the standard deviation for teacher candidates from moderately selective institutions is 25.2 while the standard deviation for teacher candidates from selective institutions is 22.3.
• On the Math subtest, the standard deviation for teacher candidates from moderately selective institutions is 33.9 while the standard deviation for teacher candidates from selective institutions is 31.9.

• On the Science subtest, the standard deviation for teacher candidates from moderately selective institutions is 31.5 while the standard deviation for teacher candidates from selective institutions is 28.5.

• On the Social Studies subtest, the standard deviation for teacher candidates from moderately selective institutions is 32.0 while the standard deviation for teacher candidates from selective institutions is 30.0.

• On the Writing subtest, the standard deviation for teacher candidates from moderately selective institutions is 30.4 while the standard deviation for teacher candidates from selective institutions is 30.8.

These results will be further addressed in Chapter Five.

Research Question 2

2. Do statewide scores on the MoGEA subtests for the 2013-2014 academic year represent a normal distribution?

Table 13 displays the results of the K-S test for normality for each of the five MoGEA subtests.

Table 13

<table>
<thead>
<tr>
<th></th>
<th>Statistic</th>
<th>Df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English &amp; Language Arts</td>
<td>.072</td>
<td>989</td>
<td>.000</td>
</tr>
<tr>
<td>Math</td>
<td>.099</td>
<td>1006</td>
<td>.000</td>
</tr>
<tr>
<td>Science</td>
<td>.053</td>
<td>990</td>
<td>.000</td>
</tr>
<tr>
<td>Social Studies</td>
<td>.063</td>
<td>987</td>
<td>.000</td>
</tr>
<tr>
<td>Writing</td>
<td>.211</td>
<td>989</td>
<td>.000</td>
</tr>
</tbody>
</table>
According to the results of the K-S test displayed in Table 13, the English and Language Arts subtest, $D(989) = 0.07, p < .001$, Math subtest, $D(1006) = 0.10, p < .001$, Science subtest, $D(990) = 0.05, p < .001$, Social Studies subtest, $D(987) = 0.06, p < .001$, and Writing subtest, $D(989) = 0.21, p < .001$, were all significantly non-normal. However, because the sample size for each subtest was quite large, histograms and normal Q-Q plots were examined for the results of each of the five subtests to determine whether the scores were truly non-normal (Field, 2009). Per the recommendation of Field (2009), Q-Q plots were used because they plot quantiles rather than all individual scores, which would be impractical due to the large sample sizes for the subtests.

Figures 19-23 display the data set for each subtest in normal Q-Q plots. Figures 4-8, presented in the analysis for RQ1, display the same data sets in histograms. (It should be noted that the MoGEA Writing subtest is scored quite differently from the other subtests. Fewer total scores are possible on this subtest, and the resulting histograms and Q-Q plots are therefore much different than the same data displays for the other subtests.)
Figure 19. Normal Q-Q Plot of MoGEA English and Language Arts subtest scores for the total population of the study.

Figure 20. Normal Q-Q Plot of MoGEA Math subtest scores for the total population of the study.
Figure 21. Normal Q-Q Plot of MoGEA Science subtest scores for the total population of the study.

Figure 22. Normal Q-Q Plot of MoGEA Social Studies subtest scores for the total population of the study.
Despite significant results from the K-S test for normality, analysis of the histograms and normal Q-Q plots for all five subtests revealed the scores on each represented normal distributions. For each subtest, specific scores became less frequent as they deviated further from the center of the data set. While the histograms (Figures 1-5) show slight deviations from symmetry—slight negative skew for English and Language Arts in Figure 1 and slight positive skew for Math in Figure 2—and the Q-Q (Figures 23-27) plots do not perfectly resemble the values expected from a normal distribution, the researcher determined their deviations from normality were not significant because of their large sample sizes. Had the histograms and Q-Q plots revealed teacher candidate subtest scores to be distributed non-normally, further statistical analyses would have been prevented due to the necessary assumption of
normally distributed data (Field, 2009). However, because the scores for each subtest represented normal distributions, further statistical analyses was conducted.

The previous data and statistical analysis, as it relates to research question two, did not support the following null hypothesis.

\( H_{01} \): Statewide scores on the MoGEA for the 2013-2014 academic year do not represent a normal distribution.

The researcher therefore rejected the null hypothesis (\( H_{01} \)).

**Research Question 3**

3. What are the mean and median scores for each decile within each MoGEA subtest?

The mean and median scores for each decile within each of the MoGEA subtests are displayed in Table 14. Each decile represents 10% of the scores for that subtest, when the scores are arranged in ascending order. Decile 1 consists of the first (lowest) 10%, decile 2 consists of the second 10%, and so on through decile 10, which consists of the final (highest) 10%.
Table 14

<table>
<thead>
<tr>
<th>Subtest Mean and Median Scores by Decile</th>
<th>English &amp; LA</th>
<th>Math</th>
<th>Science</th>
<th>Social Studies</th>
<th>Writing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decile 1</td>
<td>188, 191</td>
<td>154, 158</td>
<td>168, 171</td>
<td>166, 171</td>
<td>184, 146</td>
</tr>
<tr>
<td>Decile 2</td>
<td>205, 203</td>
<td>170, 171</td>
<td>191, 189</td>
<td>186, 183</td>
<td>193, 193</td>
</tr>
<tr>
<td>Decile 3</td>
<td>215, 214</td>
<td>179, 177</td>
<td>201, 202</td>
<td>195, 195</td>
<td>193, 193*</td>
</tr>
<tr>
<td>Decile 4</td>
<td>224, 226</td>
<td>187, 189</td>
<td>206, 208</td>
<td>205, 208</td>
<td>211, 220</td>
</tr>
<tr>
<td>Decile 5</td>
<td>230, 231</td>
<td>195, 195</td>
<td>218, 220</td>
<td>215, 214</td>
<td>220, 220*</td>
</tr>
<tr>
<td>Decile 6</td>
<td>236, 237</td>
<td>204, 202</td>
<td>227, 226</td>
<td>222, 220</td>
<td>238, 247</td>
</tr>
<tr>
<td>Decile 7</td>
<td>242, 243</td>
<td>214, 214</td>
<td>234, 232</td>
<td>232, 232</td>
<td>247, 247*</td>
</tr>
<tr>
<td>Decile 8</td>
<td>250, 249</td>
<td>225, 226</td>
<td>244, 245</td>
<td>242, 238</td>
<td>247, 247*</td>
</tr>
<tr>
<td>Decile 9</td>
<td>257, 254</td>
<td>241, 238</td>
<td>256, 257</td>
<td>253, 251</td>
<td>248, 247*</td>
</tr>
<tr>
<td>Decile 10</td>
<td>269, 266</td>
<td>267, 263</td>
<td>275, 272</td>
<td>272, 271</td>
<td>280, 273</td>
</tr>
</tbody>
</table>

* These median scores are the same as those in the previous decile because of the nature in which the Writing subtest is scored. There are fewer possible score totals for the Writing subtest, so many teacher candidates obtained the same scores.

This decile analysis echoed the results of the normality tests performed in research question 2. When the differences in mean scores between consecutive deciles for the English and Language Arts subtest were analyzed, for example, the data revealed these differences decreased as the decile number increased: 17, 10, 9, 6, 6, 8, 7, 8. Scores became less varied between deciles after the lower four deciles, indicating a negative skew to the data distribution. Likewise, for the Math subtest, scores were less varied between deciles 2 and 7 than in the upper three deciles, indicating a positive skew to the distribution: 16, 9, 8, 8, 9, 10, 11, 16, 16. The other three subtests’ mean differences between deciles are more representative of normal distributions, varying less between the center deciles.

This decile analysis also gave the researcher a general idea of the level of difficulty of the subtests in comparison with each other. The maximum score for each subtest is 300 points, but within every decile across all five subtests, the math mean and
median scores are the lowest. The Math subtest therefore appears to be the most challenging for teacher candidates. Similarly, the English and Language Arts subtest mean and median scores are the highest within every decile across all five subtests, which indicated it is the least challenging for teacher candidates. (It should be noted the Writing subtest had higher mean and median scores than the English and Language Arts subtest in deciles 6, 7, and 10; however, due to the nature in which the latter subtest is scored, decile analysis is less practical in application and was considered less significant by the researcher.)

**Research Question 4**

4. When composite ACT score is used as a covariate, is there a difference in MoGEA mean subtest scores for the following demographic variables?

   a. Ethnic origin
   
   b. Gender
   
   c. Total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained
   
   d. Cumulative grade point average at the time the student’s highest reported score on the MoGEA was obtained
   
   e. Classification of the candidate as receiving financial aid or not receiving financial aid
   
   f. Class level

As recommended by Field (2009), prior to conducting an ANCOA, a one-way ANOVA was conducted to determine whether teacher candidates’ subtest scores were
independent of the intended covariate: candidates’ composite ACT scores. Table 15 displays the results of the one-way ANOVA test performed for each of the five subtests.

Table 15

<table>
<thead>
<tr>
<th>ANOVA – ACT and Subtest Scores</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English &amp; Language Arts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>148942.04</td>
<td>22</td>
<td>6770.09</td>
<td>18.87</td>
<td>.00</td>
</tr>
<tr>
<td>Within Groups</td>
<td>246866.19</td>
<td>688</td>
<td>358.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>395808.23</td>
<td>710</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>366036.87</td>
<td>22</td>
<td>16638.04</td>
<td>26.76</td>
<td>.00</td>
</tr>
<tr>
<td>Within Groups</td>
<td>430287.35</td>
<td>692</td>
<td>621.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>796324.22</td>
<td>714</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>292224.99</td>
<td>22</td>
<td>13282.95</td>
<td>24.95</td>
<td>.00</td>
</tr>
<tr>
<td>Within Groups</td>
<td>365730.90</td>
<td>687</td>
<td>532.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>657955.89</td>
<td>709</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>280057.62</td>
<td>22</td>
<td>12729.89</td>
<td>21.48</td>
<td>.00</td>
</tr>
<tr>
<td>Within Groups</td>
<td>406481.36</td>
<td>686</td>
<td>592.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>686538.98</td>
<td>708</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>111160.87</td>
<td>22</td>
<td>5052.77</td>
<td>6.38</td>
<td>.00</td>
</tr>
<tr>
<td>Within Groups</td>
<td>547430.37</td>
<td>691</td>
<td>792.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>658591.24</td>
<td>713</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There was a significant effect of ACT composite score on the English and Language Arts subtest, $F(22, 688) = 18.87, p < .01$. There was a significant effect of ACT composite score on the Math subtest, $F(22, 692) = 26.76, p < .01$. There was a significant effect of ACT composite score on the Science subtest, $F(22, 687) = 24.95, p < .01$. There was a significant effect of ACT composite score on the Social Studies subtest, $F(22, 686) = 21.48, p < .01$. There was a significant effect of ACT composite score on the English and Language Arts subtest, $F(22, 691) = 6.38, p < .01$. The results of the one-way ANOVA test for all subtests showed MoGEA subtest scores were significantly dependent upon ACT composite scores. Therefore, the ANCOVA test was not conducted because, according to Field (2009), ACT composite score could not be used as a
covariate once shown to have a significant effect on the subtest scores. Post hoc tests could not be performed on ACT composite scores because scores ranged from 13 to 34, creating too many groups.

Although, per the recommendation of Field (2009), an ANCOVA test was not conducted, the researcher conducted additional ANOVA tests to determine MoGEA subtest scores were independent of teacher candidates’ self-reported ethnicity and class level. Table 16 displays the results of the one-way ANOVA test involving teacher candidates’ MoGEA subtest scores and their self-reported ethnicity. For the purposes of the data analysis, teacher candidates’ ethnicities were recoded as nominal numerical variables ranging from 0 through 7 and corresponding with the ethnicities identified in Table 4.

Table 16

ANOVA – Ethnicity and Subtest Scores

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English &amp; Language Arts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>8079.71</td>
<td>7</td>
<td>1154.25</td>
<td>2.01</td>
<td>.051</td>
</tr>
<tr>
<td>Within Groups</td>
<td>477565.97</td>
<td>833</td>
<td>573.308</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>485645.68</td>
<td>840</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>12076.58</td>
<td>7</td>
<td>1725.23</td>
<td>1.53</td>
<td>.153</td>
</tr>
<tr>
<td>Within Groups</td>
<td>952007.91</td>
<td>845</td>
<td>1126.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>964084.49</td>
<td>852</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>19665.88</td>
<td>7</td>
<td>2809.41</td>
<td>2.96</td>
<td>.01</td>
</tr>
<tr>
<td>Within Groups</td>
<td>789705.14</td>
<td>832</td>
<td>949.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>809371.02</td>
<td>839</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>19322.42</td>
<td>7</td>
<td>2760.35</td>
<td>2.79</td>
<td>.01</td>
</tr>
<tr>
<td>Within Groups</td>
<td>822022.31</td>
<td>831</td>
<td>989.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>841344.73</td>
<td>838</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>16076.30</td>
<td>7</td>
<td>2296.62</td>
<td>2.44</td>
<td>.02</td>
</tr>
<tr>
<td>Within Groups</td>
<td>783852.53</td>
<td>833</td>
<td>941.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>799928.83</td>
<td>840</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
There was a significant effect of ethnicity on the Science subtest, $F(7, 832) = 2.96, p < .01$. There was a significant effect of ethnicity on the Social Studies subtest, $F(7, 831) = 2.79, p < .01$. There was a significant effect of ethnicity on the Writing subtest, $F(7, 833) = 2.44, p < .05$. The results of the one-way ANOVA test for the Science, Social Studies, and Writing subtests showed MoGEA subtest scores were significantly dependent upon teacher candidate ethnicity.

Post hoc (Tukey) tests were performed on these three subtests, as recommended by Field (2009), and significant differences between ethnic groups were revealed on the Science and Social Studies subtests. The Tukey post hoc procedure was specifically chosen because it is the most powerful post hoc procedure “when testing large numbers of means” (Field, 2009, p. 374), as was the case in this study. On the Science subtest, the significant mean difference was between the Alaskan Native/American Indian and Non-Resident Alien ethnic groups. Scores from only two teacher candidates who self-identified as Non-Resident Alien were submitted; therefore, the researcher did not consider this result significant. On the Social Studies subtest, significant mean differences were revealed between the Alaskan Native/American Indian and Black, Non-Hispanic ethnic groups as well as the Alaskan Native/American Indian and the Unknown/Not Reported groups. Because the teacher candidates in the Unknown/Not Reported group did not identify or report a specific ethnicity, the post hoc (Tukey) test results for this comparison were not considered significant. The post hoc (Tukey) test performed on the Writing subtests did not reveal significant differences despite the results of the original ANOVA. Table 17 displays the results of the post hoc (Tukey) test for the significant Social Studies subtest.
Table 17

Post Hoc (Tukey) Multiple Comparisons Between Ethnic Groups – Social Studies Subtest

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Ethnicity</th>
<th>Mean Difference</th>
<th>SE</th>
<th>Sig.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaskan Native/ American Indian</td>
<td>Black, Non-Hispanic</td>
<td>32.10</td>
<td>10.53</td>
<td>.05</td>
<td>.09</td>
<td>64.11</td>
</tr>
<tr>
<td>Black, Non-Hispanic</td>
<td>Alaskan Native/ American Indian</td>
<td>-32.10</td>
<td>10.53</td>
<td>.05</td>
<td>-64.11</td>
<td>-.09</td>
</tr>
</tbody>
</table>

Table 18 displays the results of the one-way ANOVA test involving teacher candidates’ MoGEA subtest scores and their class level. For the purposes of the data analysis, teacher candidates’ class levels were recoded as nominal numerical variables ranging from 0 through 4 and corresponding with the class levels identified in Table 9.

Table 18

ANOVA – Class Level and Subtest Scores

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Between Groups</th>
<th>df</th>
<th>Within Groups</th>
<th>Total</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English &amp; Language Arts</td>
<td>2429.39</td>
<td>4</td>
<td>460806.85</td>
<td>463236.24</td>
<td>607.35</td>
<td>1.09</td>
<td>.36</td>
</tr>
<tr>
<td>Math</td>
<td>14088.80</td>
<td>4</td>
<td>844218.11</td>
<td>858306.91</td>
<td>3522.20</td>
<td>3.50</td>
<td>.01</td>
</tr>
<tr>
<td>Science</td>
<td>1936.18</td>
<td>4</td>
<td>738904.60</td>
<td>740840.77</td>
<td>484.04</td>
<td>.54</td>
<td>.70</td>
</tr>
<tr>
<td>Social Studies</td>
<td>7555.74</td>
<td>4</td>
<td>766060.08</td>
<td>773615.82</td>
<td>1888.93</td>
<td>2.03</td>
<td>.09</td>
</tr>
<tr>
<td>Writing</td>
<td>4568.07</td>
<td>4</td>
<td>739825.03</td>
<td>744393.11</td>
<td>1142.02</td>
<td>1.28</td>
<td>.27</td>
</tr>
</tbody>
</table>

The results of the one-way ANOVA showed there was a significant effect of class level on the MoGEA Math subtest, $F(4, 839) = 3.50, p < .01$. A post hoc (Tukey) test was performed on the Math subtest, and significant mean differences were revealed between
sophomore and seniors and between juniors and seniors. Table 19 displays the results of the post hoc (Tukey) test for the significant Math subtest.

Table 19

<table>
<thead>
<tr>
<th>Class Level</th>
<th>Class Level</th>
<th>Mean Difference</th>
<th>SE</th>
<th>Sig.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sophomores</td>
<td>Seniors</td>
<td>10.18</td>
<td>3.25</td>
<td>.02</td>
<td>1.29</td>
<td>19.07</td>
</tr>
<tr>
<td>Juniors</td>
<td>Seniors</td>
<td>8.88</td>
<td>3.17</td>
<td>.04</td>
<td>.23</td>
<td>17.53</td>
</tr>
<tr>
<td>Seniors</td>
<td>Sophomores</td>
<td>-10.18</td>
<td>3.25</td>
<td>.02</td>
<td>-19.07</td>
<td>-1.29</td>
</tr>
<tr>
<td>Seniors</td>
<td>Juniors</td>
<td>-8.88</td>
<td>3.17</td>
<td>.04</td>
<td>-17.53</td>
<td>-.23</td>
</tr>
</tbody>
</table>

In addition to the one-way ANOVA tests, the researcher also conducted independent-samples t-tests to determine whether significant differences in mean scores existed on the five MoGEA subtests for various groups within the overall population of the study. For ethnicity, the study group was divided based on populations that have historically been shown to be negatively affected by test bias (Gitomer et al., 2011; Angrist & Guryan, 2004). Teacher candidates who identified American Indian/Alaskan Native, Asian or Pacific Islander, Non-Resident Alien, Bi- or Multi-Racial, and White, Non-Hispanic were placed into Group 1. Teacher candidates who identified as Black, Non-Hispanic or Hispanic/Latino, both groups shown to be negatively affected by test bias, were placed into Group 2. Table 20 displays the results of the independent-samples t-test.
Table 20

*Independent-Samples t-Test – Ethnicity*

<table>
<thead>
<tr>
<th>Group</th>
<th>1</th>
<th>2</th>
<th>Paired</th>
<th>Mean</th>
<th>SD</th>
<th>SE</th>
<th>T</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English &amp; LA</td>
<td>743</td>
<td>75</td>
<td>6.50</td>
<td>233</td>
<td>23.52</td>
<td>.86</td>
<td>2.26</td>
<td>816</td>
<td>0.02</td>
</tr>
<tr>
<td>Math</td>
<td>753</td>
<td>76</td>
<td>10.83</td>
<td>205</td>
<td>33.40</td>
<td>1.22</td>
<td>2.70</td>
<td>827</td>
<td>0.01</td>
</tr>
<tr>
<td>Science</td>
<td>745</td>
<td>72</td>
<td>10.26</td>
<td>225</td>
<td>30.70</td>
<td>1.13</td>
<td>2.70</td>
<td>815</td>
<td>0.01</td>
</tr>
<tr>
<td>Social Studies</td>
<td>743</td>
<td>72</td>
<td>9.29</td>
<td>220</td>
<td>30.91</td>
<td>1.13</td>
<td>2.38</td>
<td>813</td>
<td>0.02</td>
</tr>
<tr>
<td>Writing</td>
<td>750</td>
<td>72</td>
<td>5.88</td>
<td>231</td>
<td>29.23</td>
<td>1.07</td>
<td>1.62</td>
<td>820</td>
<td>0.11</td>
</tr>
</tbody>
</table>

For the English and Language Arts subtest, on average, members of Group 1 scored higher ($M = 233, SE = .86$) than did members of Group 2 ($M = 226, SE = 2.94$), $t(816) = 2.26, p < .05$. For the Math subtest, on average, members of Group 1 scored higher ($M = 205, SE = 1.22$) than did members of Group 2 ($M = 194, SE = 3.78$), $t(827) = 2.70, p < .01$. For the Science subtest, on average, members of Group 1 scored higher ($M = 225, SE = 1.13$) than did members of Group 2 ($M = 214, SE = 3.77$), $t(815) = 2.70, p < .01$. For the Social Studies subtest, on average, members of Group 1 scored higher ($M = 220, SE = 1.13$) than did members of Group 2 ($M = 210, SE = 4.55$), $t(813) = 2.38, p < .05$. For the Writing subtest, on average, members of Group 1 scored higher ($M = 231, SE = 1.07$) than did members of Group 2 ($M = 225, SE = 3.68$), $t(820) = 1.62, p > .05$. For all but the Writing subtest, the 2-tailed probability level was significant. Additionally, on all five subtests, Group 1 scored higher than Group 2.
For gender, teacher candidates who identified as female were placed into Group 1. Teacher candidates who identified as male were placed into Group 2. Table 21 displays the results of the independent-samples t-test.

Table 21

**Independent-Samples t-Test – Gender**

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SE</th>
<th>T</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English &amp; LA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>753</td>
<td>231</td>
<td>23.73</td>
<td>.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>214</td>
<td>236</td>
<td>23.68</td>
<td>1.62</td>
<td>-2.58</td>
<td>965</td>
<td>.01</td>
</tr>
<tr>
<td>Paired</td>
<td></td>
<td>-4.73</td>
<td></td>
<td></td>
<td>-1.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>764</td>
<td>200</td>
<td>31.99</td>
<td>1.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>218</td>
<td>218</td>
<td>33.36</td>
<td>2.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired</td>
<td></td>
<td>-17.47</td>
<td></td>
<td></td>
<td>2.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>750</td>
<td>219</td>
<td>29.32</td>
<td>1.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>218</td>
<td>235</td>
<td>31.70</td>
<td>2.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired</td>
<td></td>
<td>-15.85</td>
<td></td>
<td></td>
<td>2.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>749</td>
<td>214</td>
<td>29.44</td>
<td>1.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>216</td>
<td>236</td>
<td>31.20</td>
<td>2.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired</td>
<td></td>
<td>-22.05</td>
<td></td>
<td></td>
<td>2.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>751</td>
<td>230</td>
<td>28.87</td>
<td>1.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>220</td>
<td>229</td>
<td>31.23</td>
<td>2.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired</td>
<td></td>
<td>1.73</td>
<td></td>
<td></td>
<td>2.26</td>
<td></td>
<td>.44</td>
</tr>
</tbody>
</table>

For the English and Language Arts subtest, on average, members of Group 1 scored lower \( (M = 231, SE = .86) \) than did members of Group 2 \( (M = 236, SE = 1.62) \), \( t(965) = -2.58, p < .05 \). For the Math subtest, on average, members of Group 1 scored lower \( (M = 200, SE = 1.16) \) than did members of Group 2 \( (M = 218, SE = 2.26) \), \( t(980) = -7.04, p < .01 \). For the Science subtest, on average, members of Group 1 scored lower \( (M = 219, SE = 1.07) \) than did members of Group 2 \( (M = 235, SE = 2.15) \), \( t(966) = -6.90, p < .01 \). For the Social Studies subtest, on average, members of Group 1 scored lower \( (M = 214, SE = 1.08) \) than did members of Group 2 \( (M = 236, SE = 2.12) \), \( t(963) = 2.38, p < .01 \). For the Writing subtest, on average, members of Group 1 scored higher \( (M = 230, SE = 1.05) \) than did members of Group 2 \( (M = 229, SE = 2.26) \), \( t(969) = .77, p > .05 \). For all
but the Writing subtest, the 2-tailed probability level was significant. Additionally, in all but the Writing subtest, Group 1 scored lower than Group 2.

For financial aid, teacher candidates who did not receive financial aid were placed into Group 1. Teacher candidates who received financial aid were placed into Group 2. Table 22 displays the results of the independent-samples t-test.

Table 22

<table>
<thead>
<tr>
<th>Independent-Samples t-Test – Financial Aid</th>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>SE</th>
<th>T</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English &amp; LA</td>
<td>1</td>
<td>165</td>
<td>231</td>
<td>20.62</td>
<td>1.61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>309</td>
<td>232</td>
<td>23.79</td>
<td>1.35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paired</td>
<td></td>
<td>-1.64</td>
<td>2.19</td>
<td>-.75</td>
<td>472</td>
<td>.46</td>
<td></td>
</tr>
<tr>
<td>Math</td>
<td>1</td>
<td>173</td>
<td>197</td>
<td>27.03</td>
<td>2.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>309</td>
<td>205</td>
<td>30.82</td>
<td>1.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paired</td>
<td></td>
<td>-7.98</td>
<td>2.80</td>
<td>-.85</td>
<td>480</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td>1</td>
<td>168</td>
<td>221</td>
<td>25.93</td>
<td>2.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>306</td>
<td>226</td>
<td>29.34</td>
<td>1.68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paired</td>
<td></td>
<td>-4.95</td>
<td>2.71</td>
<td>-.83</td>
<td>472</td>
<td>.07</td>
<td></td>
</tr>
<tr>
<td>Social Studies</td>
<td>1</td>
<td>165</td>
<td>217</td>
<td>27.91</td>
<td>2.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>305</td>
<td>221</td>
<td>29.68</td>
<td>1.70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paired</td>
<td></td>
<td>-4.00</td>
<td>2.81</td>
<td>-.42</td>
<td>468</td>
<td>.16</td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td>1</td>
<td>166</td>
<td>229</td>
<td>28.45</td>
<td>2.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>305</td>
<td>232</td>
<td>28.32</td>
<td>1.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paired</td>
<td></td>
<td>-2.56</td>
<td>2.74</td>
<td>-.256</td>
<td>469</td>
<td>.35</td>
<td></td>
</tr>
</tbody>
</table>

For the English and Language Arts subtest, on average, members of Group 1 scored lower ($M = 231, SE = 1.61$) than did members of Group 2 ($M = 232, SE = 2.19$), $t(472) = -.75, p > .05$. For the Math subtest, on average, members of Group 1 scored lower ($M = 197, SE = 2.06$) than did members of Group 2 ($M = 205, SE = 1.75$), $t(480) = -2.85, p < .01$. For the Science subtest, on average, members of Group 1 scored lower ($M = 221, SE = 2.00$) than did members of Group 2 ($M = 226, SE = 1.68$), $t(472) = -1.83, p > .05$. For the Social Studies subtest, on average, members of Group 1 scored lower ($M = 217, SE = 2.17$) than did members of Group 2 ($M = 221, SE = 1.70$), $t(468) = -1.42, p >
.05. For the Writing subtest, on average, members of Group 1 scored lower \((M = 229, SE = 2.21)\) than did members of Group 2 \((M = 232, SE = 1.62)\), \(t(469) = 1.62, p > .05\). The 2-tailed probability level was significant on only the Math subtest, \(p < .01\).

The previous data and statistical analysis, as it relates to Research Question Four, did not support the following null hypothesis.

\(H_02: \) There is no difference in MoGEA mean subtest scores between demographic variables.

The researcher therefore rejected the null hypothesis \((H_02)\).

**Research Question 5**

5. How will establishing cut scores at the lowest score within each decile affect teacher candidate admission for populations that have historically been shown to be negatively affected by test bias?

As was previously discussed in prior chapters of this study, tests similar to the MoGEA are often biased against teacher candidates from African American and Hispanic populations and therefore negatively affect the overall performance of test takers from these ethnic groups (Gitomer et al., 2011; Angrist & Guryan, 2004). To better examine whether the MoGEA is potentially biased toward these teacher candidate populations, the scores for each subtest were broken out into two groups. The first group consisted of Black, Non-Hispanic and Hispanic/Latino candidates. The second consisted of the other ethnic groups from the study: American Indian/Alaskan Native, Asian or Pacific Islander, non-resident alien, two or more races, and white, non-Hispanic. Scores from candidates who did not report an ethnic group were not included. Table 23 displays the passing rates (expressed as a percent) for both of these groups and the total population of the study at
ten different possible cut score marks. For each of the five MoGEA subtests, these ten cut scores were determined by identifying the minimum score reported within each decile of the set of subtest scores for the total study group population.

Table 23

*Table continued on next page.*
Table 23 (continued from previous page)

Possible MoGEA Cut Scores by Decile and Ethnic Groups

<table>
<thead>
<tr>
<th>Possible Subtest</th>
<th>Black, Non-Hispanic &amp; Hispanic/Latino Passing Rates (%)</th>
<th>Other Ethnic Groups Passing Rates (%)</th>
<th>Total Population Passing Rates (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science</td>
<td>n = 73</td>
<td>n = 749</td>
<td>n = 996</td>
</tr>
<tr>
<td>134</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>183</td>
<td>86.1</td>
<td>92.1</td>
<td>91.8</td>
</tr>
<tr>
<td>195</td>
<td>77.8</td>
<td>86.2</td>
<td>84.8</td>
</tr>
<tr>
<td>208</td>
<td>52.1</td>
<td>74.0</td>
<td>71.2</td>
</tr>
<tr>
<td>214</td>
<td>47.2</td>
<td>65.4</td>
<td>63.4</td>
</tr>
<tr>
<td>220</td>
<td>40.3</td>
<td>60.0</td>
<td>57.5</td>
</tr>
<tr>
<td>232</td>
<td>33.3</td>
<td>45.8</td>
<td>42.4</td>
</tr>
<tr>
<td>238</td>
<td>30.1</td>
<td>36.8</td>
<td>34.4</td>
</tr>
<tr>
<td>251</td>
<td>18.1</td>
<td>24.2</td>
<td>22.6</td>
</tr>
<tr>
<td>263</td>
<td>6.9</td>
<td>14.2</td>
<td>13.1</td>
</tr>
<tr>
<td>Social Studies</td>
<td>n = 75</td>
<td>n = 751</td>
<td>n = 998</td>
</tr>
<tr>
<td>134</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>183</td>
<td>76.0</td>
<td>90.9</td>
<td>90.2</td>
</tr>
<tr>
<td>189</td>
<td>66.7</td>
<td>79.4</td>
<td>85.0</td>
</tr>
<tr>
<td>202</td>
<td>56.0</td>
<td>66.3</td>
<td>71.8</td>
</tr>
<tr>
<td>208</td>
<td>53.3</td>
<td>59.4</td>
<td>65.7</td>
</tr>
<tr>
<td>220</td>
<td>49.3</td>
<td>52.7</td>
<td>52.8</td>
</tr>
<tr>
<td>226</td>
<td>40.0</td>
<td>38.6</td>
<td>45.0</td>
</tr>
<tr>
<td>238</td>
<td>37.3</td>
<td>32.2</td>
<td>32.7</td>
</tr>
<tr>
<td>247</td>
<td>17.3</td>
<td>21.8</td>
<td>22.1</td>
</tr>
<tr>
<td>263</td>
<td>12.0</td>
<td>11.7</td>
<td>11.4</td>
</tr>
<tr>
<td>Writing</td>
<td>n = 73</td>
<td>n = 753</td>
<td>n = 998</td>
</tr>
<tr>
<td>140</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>193</td>
<td>90.4</td>
<td>97.9</td>
<td>96.8</td>
</tr>
<tr>
<td>193</td>
<td>90.4</td>
<td>97.9</td>
<td>96.8</td>
</tr>
<tr>
<td>220</td>
<td>58.9</td>
<td>67.5</td>
<td>66.2</td>
</tr>
<tr>
<td>220</td>
<td>58.9</td>
<td>67.5</td>
<td>66.2</td>
</tr>
<tr>
<td>247</td>
<td>39.7</td>
<td>48.9</td>
<td>46.6</td>
</tr>
<tr>
<td>247</td>
<td>39.7</td>
<td>48.9</td>
<td>46.6</td>
</tr>
<tr>
<td>263</td>
<td>6.8</td>
<td>12.0</td>
<td>11.4</td>
</tr>
</tbody>
</table>

Because the passing rates for the Black, Non-Hispanic and Hispanic/Latino group are consistently lower than those of the other group, the following null hypothesis is not supported.
H$_0$3: Establishing cut scores at the lowest score within each decile has no effect on teacher candidate admission for populations that have historically been shown to be negatively affected by test bias.

The researcher therefore rejected the null hypothesis (H$_0$3).

Per the literature review presented in Chapter Two of the study, the differences in passing rates for these two groups reveals significant potential for bias against the Black, Non-Hispanic and Hispanic/Latino teacher candidate populations. Despite attempts made by Pearson’s Evaluation Systems Group to ensure the MoGEA subtests were not biased toward any population, including the formation of a bias review committee that reviewed and approved the MoGEA framework, the results of this analysis indicated the goal of creating an unbiased test of basic knowledge for Missouri’s teacher candidates was not achieved. Chapter Five will further address these findings.

**Research Question 6**

6. Is it possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study?

First, Pearson Product-Moment Correlation Coefficient (PPMCC) was computed to determine whether the study’s dependent variables, teacher candidates’ scores on the five MoGEA subtests, were significantly correlated with the following demographic characteristics of the study’s population: total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained, cumulative GPA at the time the student’s higher reported score on the MoGEA was obtained, and maximum composite ACT score obtained by the teacher candidate (Field, 2009). Table 24 displays the results of the PPMCC tests for each subtest.
Table 24

**PPMCC Results for MoGEA Subtest Scores and Independent Variables**

<table>
<thead>
<tr>
<th>Subtest</th>
<th>PPMCC</th>
<th>Credit Hours</th>
<th>GPA</th>
<th>ACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>English &amp; LA</td>
<td>.08</td>
<td>.28</td>
<td>.62</td>
<td></td>
</tr>
<tr>
<td>n = 989</td>
<td>.01</td>
<td>.08</td>
<td>.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.02</td>
<td>.00</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>809</td>
<td>884</td>
<td>669</td>
<td></td>
</tr>
<tr>
<td>Math</td>
<td>.04</td>
<td>.32</td>
<td>.68</td>
<td></td>
</tr>
<tr>
<td>n = 1006</td>
<td>.00</td>
<td>.10</td>
<td>.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.31</td>
<td>.00</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>817</td>
<td>898</td>
<td>672</td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td>.12</td>
<td>.26</td>
<td>.67</td>
<td></td>
</tr>
<tr>
<td>n = 990</td>
<td>.01</td>
<td>.07</td>
<td>.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>801</td>
<td>883</td>
<td>668</td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td>.17</td>
<td>.25</td>
<td>.63</td>
<td></td>
</tr>
<tr>
<td>n = 987</td>
<td>.03</td>
<td>.06</td>
<td>.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>798</td>
<td>881</td>
<td>668</td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td>-.01</td>
<td>.15</td>
<td>.24</td>
<td></td>
</tr>
<tr>
<td>n = 989</td>
<td>.00</td>
<td>.02</td>
<td>.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td>.00</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>802</td>
<td>888</td>
<td>675</td>
<td></td>
</tr>
</tbody>
</table>

Scores on the English and Language Arts subtest were significantly correlated with total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained, \( r = .08, R^2 = .01, p < .05 \), GPA at the time the student’s highest reported score on the MoGEA was obtained, \( r = .28, R^2 = .08, p < .01 \), and composite ACT score, \( r = .62, R^2 = .38, p < .01 \). Scores on the Math subtest were significantly correlated with GPA at the time the student’s highest reported score on the MoGEA was obtained, \( r = .32, R^2 = .10, p < .01 \), and composite ACT score, \( r = .68, R^2 = .46, p < .01 \). Scores on the Science subtest were significantly correlated with total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained, \( r = .12, R^2 = .01, p < .01 \), GPA at the time the student’s highest reported score
on the MoGEA was obtained, $r = .26, R^2 = .07, p < .01$, and composite ACT score, $r = .67, R^2 = .45, p < .01$. Scores on the Social Studies subtest were significantly correlated with total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained, $r = .17, R^2 = .03, p < .01$, GPA at the time the student’s highest reported score on the MoGEA was obtained, $r = .25, R^2 = .06, p < .01$, and composite ACT score, $r = .63, R^2 = .40, p < .01$. Scores on the Writing subtest were significantly correlated with GPA at the time the student’s highest reported score on the MoGEA was obtained, $r = .15, R^2 = .02, p < .01$, and composite ACT score, $r = .24, R^2 = .06, p < .01$.

Following the PPMCC test, a MANOVA was performed to determine whether any of the study’s following independent variables had an effect on teacher candidate performance on the MoGEA subtests: ethnicity, gender, total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained, GPA at the time the student’s highest reported score on the MoGEA was obtained, maximum composite ACT score obtained by the teacher candidate, and whether or not the teacher candidate received financial aid. As in previous analyses, teacher candidates were placed divided into the same two groups for ethnicity, gender, and financial aid. Table 25 displays the MANOVA results.
Table 25

**MANOVA – Independent Variables**

<table>
<thead>
<tr>
<th>Effect</th>
<th>Value</th>
<th>Effect</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnicity</td>
<td>.01</td>
<td>2.04</td>
<td>5</td>
<td>792</td>
<td>.07</td>
</tr>
<tr>
<td>Gender</td>
<td>.14</td>
<td>26.10</td>
<td>5</td>
<td>941</td>
<td>.00</td>
</tr>
<tr>
<td>Credit Hours</td>
<td>.39</td>
<td>1.60</td>
<td>153</td>
<td>637</td>
<td>.00</td>
</tr>
<tr>
<td>GPA</td>
<td>.59</td>
<td>2.39</td>
<td>172</td>
<td>691</td>
<td>.00</td>
</tr>
<tr>
<td>ACT</td>
<td>2.03</td>
<td>61.41</td>
<td>21</td>
<td>634</td>
<td>.00</td>
</tr>
<tr>
<td>Financial Aid</td>
<td>.02</td>
<td>1.59</td>
<td>5</td>
<td>462</td>
<td>.16</td>
</tr>
</tbody>
</table>

For the reporting of the above statistical analyses, the researcher chose Roy’s Largest Root to determine the overall significance because it “represents the maximum possible between-group differences given the data collected” and “should in many cases be the most powerful” MANOVA test statistic (Field, 2009, p. 603). There was a significant effect of gender on MoGEA subtest scores, $F(5, 941) = 26.10, p < .01$. There was a significant effect of total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained on MoGEA subtest scores, $F(153, 637) = 1.60, p < .01$. There was a significant effect of GPA at the time the student’s highest reported score on the MoGEA was obtained on MoGEA subtest scores, $F(172, 691) = 2.39, p < .01$. There was a significant effect of maximum composite ACT score on MoGEA subtest scores, $F(21, 634) = 61.41, p < .01$. Ethnicity and financial aid did not show significant effects on MoGEA subtests scores, $p > .05$.

Following the MANOVA, per the recommendation of Field (2009), the researcher conducted discriminant functional analysis to determine whether teacher candidates’ MoGEA subtest scores can be used to predict their gender, ethnicity, or whether or not they receive financial aid. As in previous analyses, teacher candidates were placed divided into the same two groups for ethnicity, gender, and financial aid. Table 26
displays the results of the discriminant functional analysis for these three demographic characteristics.

Table 26

<table>
<thead>
<tr>
<th>Subtest</th>
<th>% of Variance</th>
<th>$r$</th>
<th>$R^2$</th>
<th>$\Lambda$</th>
<th>df</th>
<th>$\chi^2$</th>
<th>Sig.</th>
<th>% Classified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnicity</td>
<td>100%</td>
<td>.11</td>
<td>.99</td>
<td>.99</td>
<td>5</td>
<td>10.58</td>
<td>.06</td>
<td>91.1</td>
</tr>
<tr>
<td>Math</td>
<td>.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td>.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>100%</td>
<td>.36</td>
<td>.87</td>
<td>.87</td>
<td>5</td>
<td>128</td>
<td>.00</td>
<td>78.6</td>
</tr>
<tr>
<td>Social Studies</td>
<td>.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td>.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Math</td>
<td>.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial Aid</td>
<td>100%</td>
<td>.13</td>
<td>.98</td>
<td>.98</td>
<td>5</td>
<td>7.67</td>
<td>.18</td>
<td>64.3</td>
</tr>
<tr>
<td>Math</td>
<td>.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td>.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Studies</td>
<td>.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The discriminant functional analysis for ethnicity revealed one discriminant function. This function explained 100% of the variance, canonical $R^2 = .01$. This discriminant function did not significantly differentiate the ethnic groups, $r = .11$, $\Lambda = .99$, $\chi^2(5) = 10.58$, $p > .05$. The correlations between outcomes and the discriminant function revealed that scores on the Math and Writing subtests loaded significantly on the function ($r = -.53$ for Math and $r = .46$ for Writing). Regarding these two grouping for ethnicity, 91.1% of the teacher candidates were classified correctly by the model.

The discriminant functional analysis for gender revealed one discriminant function. This function explained 100% of the variance, canonical $R^2 = .13$. This discriminant function significantly differentiated the gender groups, $r = .36$, $\Lambda = .87$, $\chi^2(5) = 128.00$, $p < .01$. The correlations between outcomes and the discriminant function revealed that scores on the Social Studies, Science, and Math subtests loaded significantly on the function ($r = .82$ for Socials Studies, $r = .61$ for Science, and $r = .59$
for Math). Regarding teacher candidates’ gender, 78.6% of the teacher candidates were classified correctly by the model.

The discriminant functional analysis for whether or not a teacher candidate received financial aid revealed one discriminant function. This function explained 100% of the variance, canonical $R^2 = .02$. This discriminant function did not significantly differentiate the financial aid groups, $r = .13$, $\Lambda = .98$, $\chi^2(5) = 7.67$, $p > .05$. The correlations between outcomes and the discriminant function revealed that scores on the Math, Science, and Social Studies subtests loaded significantly on the function ($r = .98$ for Math, $r = .62$ for Science, and $r = .52$ for Social Studies). Regarding whether or not teacher candidates received financial aid, 64.3% of the teacher candidates were classified correctly by the model.

The previous data and statistical analysis, as it relates to Research Question Six, did not support the following null hypothesis.

$H_04$: It is not possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study.

The researcher therefore rejected the null hypothesis ($H_04$).

**Summary**

The purpose of this quantitative study was to recommend statewide cut score norms for the MoGEA to the executive board of the Missouri Association of Colleges for Teacher Education (MACTE) for the purpose of informing its representatives to the Missouri Advisory Council of Certification for Educators (MACCE). Additionally, this study aimed to identify demographic characteristics of EPP candidates that might serve as
predictors of success on the MoGEA. The conceptual underpinning of this study, the process of systemic change, served to frame the study and provide a lens through which the statistical analyses were conducted and interpreted.

The statistical tests used to answer the research questions designed in accordance with this study’s purpose included descriptive summary statistics—mean, median, mode, and standard deviation—for each MoGEA subtest score for the total population of the study group and broken out by institutional selectivity (RQ1). For RQ2, a Kolmogorov-Smirnov test for normality was conducted for each subtest for the total population of the study group and broken out by institutional selectivity. Histograms and normal Q-Q plots were also analyzed for each of these normality tests because their sample sizes were quite large (Field, 2009). The researcher also performed decile analysis in order to determine the mean and median scores for each decile within each MoGEA subtest (RQ3).

For RQ4, the researcher conducted a one-way ANOVA and determined teacher candidates’ subtest scores were highly dependent upon their maximum composite ACT score. Because of this finding, an ANCOVA was not performed. However, the researcher performed separate one-way ANOVA tests for teacher candidates’ subtest scores and both ethnicity and gender to determine whether subtest scores were significantly dependent on either of these demographic characteristics. Also as part of RQ4, the researcher performed independent-samples t-tests to compare subtest scores between members of different demographic subgroups (Field, 2009), including ethnicity, gender, and whether or not teacher candidates received financial aid. Descriptive analysis was performed to answer RQ5 to determine whether teacher candidates who identify as Black, Non-Hispanic or Hispanic/Latino would pass at different rates than candidates of
other ethnicities if cut scores are set at the lowest obtained score in each decile of the total population.

Finally, in order to determine whether demographic characteristics can be used to predict success on the MoGEA (RQ6), Pearson Product-Moment Correlation Coefficient was used to examine the relationship between MoGEA subtest scores and various demographic characteristics. After the PPMCC, a MANOVA was conducted and was followed by discriminant functional analysis. These statistical analyses and their results led the researcher to fail to accept all null hypotheses associated with the study’s research questions. Chapter Five will discuss the findings, conclusions, and implications of these results.
CHAPTER FIVE

FINDINGS, CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

The decision to use the Missouri General Education Assessment (MoGEA) rather than the College Basic Academic Subjects Examination (C-BASE) is a policy change that affects all stakeholders throughout Missouri’s educational system. Because the MoGEA was a new examination and there were no prior student results on which to base cut scores for the 2013-2014 school year, educator preparation programs (EPPs) were individually responsible for setting their own cut scores for the first two years of the exam’s implementation. The availability of 2013-2014 MoGEA results afforded the researcher the opportunity to better inform the Missouri Department of Elementary and Secondary Education (DESE) and EPP administrators as they planned to set statewide subtest cut scores. Statistical analysis of 2013-2014 MoGEA results also allowed the researcher to determine whether any student demographic characteristics can be used to predict success on the MoGEA.

This chapter provides a summary of the study by reviewing the problem, purpose, research questions, and null hypotheses. As part of this overall summary, brief summaries of the literature reviewed for the study and the study group are provided. Next, a review of the findings from the statistical analyses is included. Finally, this chapter presents the conclusions and implications of the study as well as recommendations for future research.

**Summary of the Study**

The ultimate goal of this study was to analyze teacher candidate scores on the five subtests of the MoGEA for the 2013-2014 school year and recommend qualifying cut scores for each subtest. This study aimed to consolidate MoGEA subtest scores of teacher
candidates throughout the state of Missouri to gain a broader view of teacher candidate performance on this assessment in its first year of use. Additionally, it was a goal of this study to determine if specific demographic characteristics—ethnicity, gender, total credit hours, GPA, and whether the teacher candidate received financial aid—could be used to predict student success on the MoGEA subtests, based on student demographic data provided by each EPP that self-selected to be a part of this research project.

**Problem Statement**

At the time of this study’s proposal, no statewide qualifying scores existed for the Missouri General Education Assessment. In order for DESE to establish examination cut scores based on the input of Missouri’s EPP personnel and inform the final decision made by Missouri’s State Board of Education (BOE), it was critical that statewide results of the 2013-2014 MoGEA be collected and analyzed to ensure a more representative population of test scores be used to establish valid and reliable cut scores for use in 2015.

**Purpose of the Study**

The purpose of this quantitative study was to recommend statewide cut score norms for the MoGEA to the executive board of the Missouri Association of Colleges for Teacher Education (MACTE) for the purpose of informing its representatives to the Missouri Advisory Council of Certification for Educators (MACCE). Additionally, this study aimed to identify demographic characteristics of EPP candidates that might serve as predictors of success on the MoGEA. If predictor variables are determined, EPP administrators can identify which of their education students will most likely require additional support and preparation for the MoGEA. This information will also make
DESE and the BOE aware of which EPP candidate demographics are likely to be denied EPP admission based on their MoGEA scores.

**Research Questions and Null Hypotheses**

This study was designed to address the following research questions and their accompanying null hypotheses:

1. What are the descriptive summary statistics for each MoGEA subtest for the total population of the study group and when broken out by institutional selectivity?
2. Do statewide scores on the MoGEA subtests for the 2013-2014 academic year represent a normal distribution?
   \[ H_0: \text{Statewide scores on the MoGEA for the 2013-2014 academic year do not represent a normal distribution.} \]
3. What are the mean and median scores for each decile within each MoGEA subtest?
4. When composite ACT score is used as a covariate, is there a difference in MoGEA mean subtest scores for the following demographic variables?
   a. Ethnic origin
   b. Gender
   c. Total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained
   d. Cumulative grade point average at the time the student’s highest reported score on the MoGEA was obtained
   e. Classification of the candidate as receiving financial aid or not receiving financial aid
H₀2: There is no difference in MoGEA mean subtest scores between demographic variables.

5. How will establishing cut scores at the lowest score within each decile affect teacher candidate admission for populations that have historically been shown to be negatively affected by test bias?

H₀3: Establishing cut scores at the lowest score within each decile has no effect on teacher candidate admission for populations that have historically been shown to be negatively affected by test bias.

6. Is it possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study?

H₀4: It is not possible to predict whether a student will achieve qualifying scores on the MoGEA based on the independent variables identified for the study.

Review of Related Literature

For the purposes of this study, the review of literature first detailed the history of teacher certification in the United States by highlighting historical events and legislative actions that influenced licensure requirements. Second, and of particular importance to this study’s research questions, the use of basic knowledge tests as a measure of accountability for teacher quality and certification was explored. Emerging from the work of authors such as Goldhaber (2007), Blackford et al. (2012), and Gitomer et al. (2011), research regarding the history of basic knowledge tests for teachers, the advantages and disadvantages of requiring teacher candidates to pass such tests, and their effectiveness as predictors of teacher success in the classroom were discussed. Third, the issue of test bias
as it relates to minority EPP candidates, as addressed by Wakefield (2006), Gitomer et al. (2011), and Memory et al. (2003), was discussed. These three topics were then examined through the lens of systemic change, the process that underpins the conceptual framework for this study.

**Study Group**

The subjects included in the study group for this project were EPP candidates throughout the state of Missouri who had taken the MoGEA between September 1, 2013, and September 1, 2014. MoGEA subtest scores were obtained from 10 EPPs throughout Missouri creating a sample size of 1037 EPP candidates. As of September 1, 2014, 4996 students had taken the MoGEA (G. Hairston, personal communication, September 18, 2014). With a total population of 4996 at a 95% confidence level and a confidence interval of 5%, a sample size of 357 would be necessary to accurately represent the population (*Sample Size Calculator*, n.d.). Therefore, the sample size of 1037 scores collected for this study can be considered an accurate representation of the population (Field, 2009).

**Findings**

The statistical analyses used for each research questions included descriptive summary statistics for research question one (RQ1), a Kolmogorov-Smirnov (K-S) distribution test and analysis of data displays for research question two (RQ2), examination of the mean and median of each decile for research question three (RQ3), one-way analysis of variance (ANOVA) and independent-samples *t*-tests for research question four (RQ4), general descriptive analysis for research question five (RQ5), and Pearson Product-Moment Correlation Coefficient (PPMCC), multivariate analysis of
variance (MANOVA), and discriminant function analysis for research question six (RQ6). The researcher proposed a one-way analysis of covariance (ANCOVA) for RQ4, but the results of the preliminary ANOVA proved it unnecessary. The researcher therefore performed additional ANOVAs and independent-samples t-tests to further explore the research question.

**Research Question 1**

The descriptive summary statistics—mean, median, mode, and standard deviation—for each MoGEA subtest for the total population of the study and when broken out by institutional selectivity are presented in Chapter Four of the study. An analysis of the summary statistics for the total population reveals the Math subtest has the lowest mean (204), the lowest median (202), and the highest standard deviation (33.0) of all five subtests. Analysis shows the English and Language Arts subtest has the highest mean (232), the highest median (231), and the lowest standard deviation (23.8) of all five subtests. Because both subtests are scored out of a maximum of 300 points, these findings indicated the English and Language Arts subtest and the Math subtest are the least and most challenging, respectively, for teacher candidates.

A comparison of the summary statistics for moderately selective and selective institutions reveals the mean and median scores from moderately selective institutions are lower on all five subtests than those from selective institutions. On average, the mean scores from moderately selective institutions are 5.4 points lower than those from selective institutions. The Math subtest has the least difference in mean scores (2 points), and the Science subtest has the greatest difference in mean scores (8 points). Additionally, on all but the Writing subtest, the standard deviations for scores from
moderately selective institutions are higher than those from selective institutions. These findings indicated teacher candidates from selective institutions are performing better on the MoGEA than those from moderately selective institutions.

**Research Question 2**

The results of the K-S test for normality performed for each of the five MoGEA subtests for the total population of the study group are presented in Chapter Four along with their accompanying histograms and Q-Q plots. While the K-S test revealed significant scores on each subtest ($p < .01$), the histograms and Q-Q plots were analyzed due to the large sample sizes involved. These data displays revealed teacher candidate scores on all five subtests represent normal distributions, decreasing in frequency as the scores move away from the centers of the data sets. There was a slight negative skew for the English and Language Arts subtest as scores were clustered above the mean of the data set. There was also a slight positive skew for the Math subtest as scores were clustered below the mean of the data set. However, because of the large size of both data sets, neither skew was great enough to be considered significant by the researcher, as was recommended by Field (2009). The null hypothesis ($H_0$) developed for RQ2 was rejected.

**Research Question 3**

The results of the decile analysis performed for research question three are presented in Chapter Four of the study. This analysis provided further support for the findings of the statistical analyses performed in response to the first and second research questions.
• RQ1: What are the descriptive summary statistics for each MoGEA subtest for the total population of the study group and when broken out by institutional selectivity?

• RQ2: Do statewide scores on the MoGEA subtests for the 2013-2014 academic year represent a normal distribution?

The analysis also provided the researcher with an understanding of how the total population of the study group was performing on each subtest when broken out by percentiles. Decile analysis further revealed the slight negative skew for the English and Language Arts subtest and the slight positive skew for the Math subtest. Additionally, decile analysis revealed differences in overall performance on the five subtests as scores on the Math subtest were consistently lower across the deciles than scores on the other four subtests. Likewise, scores on the English and Language Arts subtests were shown to be higher across all deciles than scores on the other four subtests.

Research Question 4

As previously discussed, the researcher originally proposed an ANCOVA to determine whether significant differences in mean subtest scores between different demographic groups existed when composite ACT score was used as a covariate. However, a preliminary one-way ANOVA revealed MoGEA subtest scores are significantly dependent on teacher candidate composite ACT scores \((p < .01)\), which eliminated the possibility of using it as a covariate (Field, 2009). The result of the one-way ANOVA indicated success on the MoGEA is directly related to composite ACT score: the higher a teacher candidate’s composite ACT score, the higher a teacher candidate’s MoGEA subtest scores. For the ACT and the MoGEA, performance on one
standardized test predicts performance on the next, a finding supported by quantitative studies conducted by other researchers as well (Wall, 2008).

To further explore RQ4, the researcher also conducted a one-way ANOVA test for ethnicity. These results are presented in Chapter Four of the study. This test revealed teacher candidates’ scores on three subtests were significantly dependent upon teacher candidate ethnicity: Science ($p < .01$), Social Studies ($p < .01$), and Writing ($p < .05$). Post hoc (Tukey) results for each subtest further revealed a significant difference in mean score on the Social Studies subtest between the Alaskan Native/American Indian and Black, Non-Hispanic ethnic groups ($p < .05$). In this case, Alaskan Native/American Indian teacher candidates performed significantly better than Black, Non-Hispanic teacher candidates. The researcher determined other significant results from the ANOVA and post hoc tests were actually insignificant due to small sample sizes (scores were submitted for only two Non-Resident Alien teacher candidates) and group classification (Unknown/Not Reported).

In addition to the one-way ANOVA test for Ethnicity, the researcher also conducted a one-way ANOVA test for teacher candidate class level. These results are presented in Chapter Four of the study. This test revealed teacher candidates’ subtest scores for Math were significantly dependent upon class level ($p < .01$). Post hoc (Tukey) results for the Math subtest further revealed significant differences in mean scores between sophomores and seniors ($p < .05$) and between juniors and seniors ($p < .05$). In both cases, teacher candidates in the senior class performed significantly better than their peers in the other two classes.
Following the three one-way ANOVA tests, the researcher also conducted three independent-samples t-tests to compare mean scores on MoGEA subtests between various groups within the total population of the study group. The first, a comparison of the subtest scores of teacher candidates who self-identified as either Black, Non-Hispanic or Hispanic/Latino and those from the study’s other ethnic groups (Unknown/Not Reported was not included), revealed that, on average, Black, Non-Hispanic teacher candidates—those historically shown to be negatively affected by test bias—scored lower than those from the study’s other ethnic groups on all five subtests and would therefore be eliminated from EPPs at a higher rate should statewide cut scores be set. Two-tailed probability levels were statistically significant on four of the five subtests: English and Language Arts ($p < .05$), Math ($p < .01$), Science ($p < .01$), and Social Studies ($p < .05$).

The second independent-samples t-test, a comparison of the subtest scores of female and male teacher candidates, revealed that, on average, female teacher candidates scored lower on four of the five MoGEA subtests, each having a significant two-tailed probability level: English and Language Arts ($p < .05$), Math ($p < .01$), Science ($p < .01$), and Social Studies ($p < .01$). Female candidates would therefore be eliminated from EPPs at a higher rate than their male counterparts should statewide cut scores be set. Only on the Writing subtest did female teacher candidates score higher than their male counterparts, and the two-tailed probability level was not statistically significant for this subtest ($p = .44$).

The final independent-samples t-test, a comparison of the subtest scores of teacher candidates who received financial aid those who did not, revealed that, on average, teacher candidates who received financial aid scored higher on all five MoGEA subtests
than did teacher candidates who did not receive financial aid. It should be noted, however, that the 2-tailed probability level was only significant for the Math subtest \( p < .01 \).

The findings resulting from the one-way ANOVAs and independent-samples \( t \)-tests indicated there are differences in overall performance on the MoGEA by members of different demographic groups. Specifically, a teacher candidate’s composite ACT score, ethnicity, class level, gender, and whether or not she or he received financial aid proved to be factors in how she or he performed on the MoGEA. The null hypothesis \( (H_0) \) developed for RQ4 was therefore rejected due to the results of these statistical analyses.

**Research Question 5**

To better examine whether the MoGEA is potentially biased toward teacher candidate populations historically shown to be negatively affected by test bias, the scores for each subtest were broken out into two groups. The first group consisted of Black, Non-Hispanic and Hispanic/Latino candidates. The second consisted of the other ethnic groups from the study: American Indian/Alaskan Native, Asian or Pacific Islander, non-resident alien, two or more races, and white, non-Hispanic. (Unknown/Not Reported teacher candidate scores were not included.) The passing rates for both of these groups and the total population of the study at ten different possible cut score marks were presented in Chapter Four of the study. For each of the five MoGEA subtests, these ten cut scores were determined by identifying the minimum score reported within each decile of the set of subtest scores for the total study group population.
The passing rates for teacher candidates who self-identified as Black, Non-Hispanic or Hispanic/Latino were consistently lower on all five MoGEA subtests, regardless of decile, than were the scores of teacher candidates in the other ethnic groups. Only on the Social Studies subtest was the passing rate for Black, Non-Hispanic or Hispanic/Latino teacher candidates higher than the passing rate for teacher candidates from the study’s other ethnic groups; this occurred in the seventh, eighth, and tenth deciles of the Social Studies subtest. The differences in passing rates for these two groups were as great as 11.4% in the fifth decile of the English and Language Arts subtest, 14.7% in the seventh decile of the Math subtest, 21.9% in the fourth decile of the Science subtest, 14.9% in the second decile of the Social Studies subtest, and 9.2% in the seventh, eighth, and ninth deciles of the Writing subtest. These findings indicated a significant difference in overall performance on the MoGEA between these two ethnic groups, revealing the exam is biased against teacher candidates who self-identify as Black, Non-Hispanic or Hispanic/Latino. Teacher candidates from these two ethnic groups would be eliminated from EPPs at a higher rate than candidates from the other ethnic groups should statewide cut scores be set. The null hypothesis ($H_0$) developed for RQ5 was rejected due to the results of these statistical analyses.

**Research Question 6**

The researcher first computed a Pearson Product-Moment Correlation Coefficient (PPMCC), as recommended by Field (2009), to determine whether the study’s dependent variables, teacher candidates’ scores on the five MoGEA subtests, were significantly correlated with the following demographic characteristics of the study’s population: total number of credit hours earned at the time the student’s highest reported score on the
MoGEA was obtained, GPA at the time the student’s highest reported score on the MoGEA was obtained, and maximum composite ACT score obtained by the teacher candidate (Field, 2009). The results of the PPMCC were presented in Chapter Four of the study.

For total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained, analysis revealed significant positive correlations on three subtests: English and Language Arts (\( p < .05 \)), Science (\( p < .01 \)), and Social Studies (\( p < .01 \)). For GPA at the time the student’s highest reported score on the MoGEA was obtained, analysis revealed significant positive correlations on all five subtests (\( p < .01 \)). For maximum composite ACT score obtained by the teacher candidate, analysis revealed significant positive correlations on all five subtests (\( p < .01 \)). The positive correlations indicate that students with more credit hours, higher GPAs, and higher ACT scores perform better on specific MoGEA subtests.

Next, the researcher performed a MANOVA to determine whether any of the study’s following independent variables had an effect on teacher candidate performance on the MoGEA subtests: ethnicity, gender, total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained, GPA at the time the student’s highest reported score on the MoGEA was obtained, maximum composite ACT score obtained by the teacher candidate, and whether or not the teacher candidate received financial aid. The results of this test were presented in Chapter Four of the study. This analysis revealed the following independent variables had a significant effect on teacher candidates’ MoGEA subtest scores: gender (\( p < .01 \)), total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained
(p < .01), GPA at the time the student’s highest reported score on the MoGEA was obtained (p < .01), and maximum composite ACT score obtained by the teacher candidate (p < .01).

Following the MANOVA, per the recommendation of Field (2009), the researcher conducted discriminant functional analysis to determine whether teacher candidates’ MoGEA subtest scores can be used to predict their gender, ethnicity, or whether or not they receive financial aid. The findings of this analysis were presented in Chapter Four of the study. The analysis revealed the discriminant function significantly differentiated the gender groups (p < .01) and the Social Studies (r = .82), Science (r = .61), and Math (r = .59) subtests loaded significantly on the function. Additionally, the model correctly classified 78.6% of the teacher candidates, further supporting gender as a strong predictor variable for determining teacher candidate success on the MoGEA. Ultimately, the findings from the PPMCC, MANOVA, and discriminant functional analysis indicated several of the demographic characteristics used as independent variables in the study have predictive value in determining student success on the MoGEA. The null hypothesis (H04) developed for RQ6 was rejected due to the results of these statistical analyses.

Conclusions

The purpose of this quantitative study was to recommend statewide cut score norms for the MoGEA to the executive board of MACTE for the purpose of informing its representatives to MACCE. However, based on the results of the statistical analyses conducted in response to the study’s research questions, the researcher concludes the MoGEA should not be used in order to determine whether teacher candidates are admitted to EPPs. The scores teacher candidates receive on this test of basic knowledge
should not determine EPP acceptance because the findings indicated the MoGEA is biased against female teacher candidates and candidates who identify as Black, Non-Hispanic or Hispanic/Latino. Further conclusions of the study are explained by research question.

**Research Questions 1, 2, and 3**

Relative to RQ1, RQ2, and RQ3, the results of the statistical analyses indicated teacher candidates performed strongest on the English and Language Arts subtest and weakest on the Math subtest. In connection, these findings show the same cut score cannot be used for each subtest without disproportionately affecting subtest passing rates. Still, when the subtests were analyzed separately, the findings showed teacher candidates are earning scores on the five MoGEA subtests that correspond with those expected from a normally distributed data set. The normal distributions indicated teacher candidate scores on each subtest were predominantly consistent and that there were not too many low or high scores affecting the overall distributions. Also relative to RQ1 and of importance to the purposes of the study, the findings indicated that teacher candidates from selective institutions performed better than their counterparts from moderately selective institutions on all five MoGEA subtests.

**Research Question 4**

Relative to RQ4, the results of the statistical analyses indicated there are significant differences in MoGEA mean subtest scores amongst different sub-populations of the study group. The results of the one-way ANOVA test performed for ethnicity indicated a test bias against one or more ethnic groups, specifically teacher candidates who self-identify as Alaskan Native/American Indian and Black, Non-Hispanic. The
results of the one-way ANOVA test performed for class level indicated significant Math
subtest mean score differences between sophomores and seniors and between juniors and
seniors. Seniors performed significantly better than both these groups on the Math
subtest, indicating a direct relationship between total credit hours obtained by teacher
candidates and MoGEA Math subtest scores.

Also relative to RQ4, the results of the independent-samples $t$-test comparing
MoGEA subtest scores of the total population of the study broken down into two ethnic
groups indicated test bias against teacher candidates who identify as Black, Non-Hispanic
or Hispanic/Latino, confirming the previous findings of Gitomer et al. (2011) and Angrist
and Guryan (2004). The results of the independent-samples $t$-test comparing the MoGEA
subtest scores of the total population of the study broken out by gender also indicated test
bias against female teacher candidates. Additionally, the results of the independent-
samples $t$-test comparing the MoGEA subtest scores of the total population of the study
broken out by whether or not they received financial aid indicated students who do not
receive financial aid do not perform as well on the MoGEA subtests. The results of the
one-way ANOVA tests for ethnicity and class level and the results of these independent-
samples $t$-tests have important implications for EPP administrators, including the need for
thoughtful curricular design for students of education and the provision of test
preparation and coursework supplementation for candidates whose scores are predicted to
be lower on the MOGEA. Such measures might better support teacher candidates
throughout the educational experience and help prepare them for success on the MoGEA.

A final conclusion resulting from the statistical analyses conducted in response to
RQ4 involves teacher candidates’ composite ACT scores. The results of the one-way
ANOVA test performed for composite ACT score indicated teacher candidates’ MoGEA subtest scores are highly dependent upon their composite ACT scores. This finding indicated the predictive value of composite ACT score when determining how teacher candidates will perform on the MoGEA subtests.

**Research Question 5**

Relative to RQ5, examination of passing rates for the two population subgroups used in the decile analysis revealed teacher candidates who self-identify as Black, Non-Hispanic and Hispanic/Latino would experience greater rates of exclusion from EPPs than would their counterparts from the other ethnic groups examined in the study. Black, Non-Hispanic and Hispanic/Latino teacher candidates would be denied admission to EPPs at higher rates than candidates from the other subgroup based on every possible cut score examined for four of the five subtests: English and Language Arts, Math, Science, and Writing. Even on the Social Studies subtest, this subgroup would still be denied admission at a higher rate than their counterparts at seven of the ten cut score marks examined. The results of the decile analysis performed in response to RQ5 reveal setting statewide MoGEA subtest cut scores, regardless of where they are set, will deny teacher candidates belonging to these two ethnic groups admission to EPPs at higher rates than their peers, further confirming the research of Memore et al. (2003), Gitomer et al. (2011), Wakefield (2008), and Angrist and Guryan (2004). As the MoGEA exam currently exists, it is biased against teacher candidates who self-identify as Black, Non-Hispanic or Latino/Hispanic. This finding has important implications for EPP personnel, the executive board of MACTE, and all stakeholders in the process of teacher credentialing in the state of Missouri.
Research Question 6

Additionally, this study aimed to identify demographic characteristics of EPP candidates that might serve as predictors of success on the MoGEA. Relative to RQ6, the results of the statistical analyses indicated several of the demographic characteristics used as independent variables for the study could be used by EPP administrators and instructors to predict whether a student will achieve qualifying scores on the MoGEA. The results of the PPMCC indicated the following independent variables used in the study have predictive value in determining how teacher candidates will perform on the MoGEA subtests: total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained, GPA at the time the student’s highest reported score on the MoGEA was obtained, and maximum composite ACT score obtained by the teacher candidate.

Also relative to RQ6, the results of the MANOVA further indicated the following independent variables used in the study have predictive value in determining how teacher candidates will perform on the MoGEA subtests: gender, total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained, GPA at the time the student’s highest reported score on the MoGEA was obtained, and maximum composite ACT score obtained by the teacher candidate. The results of the PPMCC and MANOVA revealed EPP faculty and administrators can expect students with fewer credit hours, lower GPAs, and/or lower ACT scores to pass the MoGEA at lower rates than their peers.

A final conclusion resulting from the statistical analyses conducted in response to RQ6 involves teacher candidates’ gender. The results of the discriminant functional
analysis indicated a teacher candidate’s gender has predictive value in determining how she or he will perform on the MoGEA subtests. EPP faculty and administrators can expect female teacher candidates to pass the MoGEA at lower rates than their male counterparts. These findings have important implications for EPP administrators and instructors as they work to support teacher candidates throughout the educational experience and help prepare them for success on the MoGEA.

**Implications**

The findings and conclusions of the study have raised several issues and practical implications for stakeholders in the process of teacher credentialing in the state of Missouri. The primary purpose of this quantitative study was to recommend statewide cut score norms for the MoGEA to the executive board of MACTE. The information provided to the executive board would better inform its representatives to MACCE of the potential effects of establishing various statewide cut score norms. Secondly, the study aimed to identify whether any demographic characteristics of the total population of the study’s teacher candidates can be used by EPP personnel to predict success on the MoGEA. In relation to these two purposes and based upon the study’s findings and conclusions, the following issues and implications were identified: test bias against specific teacher candidate demographic groups and the predictive value of teacher candidates’ composite ACT scores and other demographic characteristics.

**MoGEA and Test Bias**

The findings and conclusions of this study, as they are supported by the data analyses selected in response to the study’s research questions, clearly indicated the MoGEA is biased against teacher candidates who self-identify as Black, Non-Hispanic or
Hispanic/Latino. As was discussed in Chapter Two of the study, much research indicates a nationwide need for an increase in the number of teachers from these ethnic groups (Gitomer et al., 2011). Wakefield (2008) foreshadowed that with an increase in high-stakes qualifying examinations, “minorities can expect to see fewer teachers of their own race teaching their children” (p. 386). The MoGEA will actually help bring his prediction to fruition in the state of Missouri as Black and Hispanic teacher candidates are prevented from becoming educators while their counterparts from other ethnic groups achieve passing MoGEA scores and enter Missouri’s classrooms as state-certified teachers.

Additionally, the mean scores of female teacher candidates were significantly lower than those of male teacher candidates on four of the five subtests, a finding that indicated test bias against females as well. It is the recommendation of the researcher that, due to its bias against Black, Hispanic, and female teacher candidates, the MoGEA as it currently exists should not be used as a qualifying exam for EPP acceptance. The MoGEA should be further reviewed, developed, and field-tested by Pearson’s Evaluation Systems Group prior to the establishment of statewide cut scores to ensure these teacher candidates are not denied admission to EPPs at higher rates than others. As it stands, should statewide cut score norms be established for the MoGEA, female teacher candidates and candidates who self-identify as Black, Non-Hispanic or Hispanic/Latino—groups already underrepresented in the teaching profession—will be disproportionately kept from joining the teaching ranks in the state of Missouri.

**Predictor Variables for MoGEA Success**

The findings and conclusions of this study, as they are supported by the data analyses selected in response to the study’s research questions, also indicated several
demographic characteristics of teacher candidates can be used by EPP administrators and instructors to predict student success on the MoGEA. The strongest predictor of teacher candidate success on the MoGEA is composite ACT score. In the statistical analyses performed in response to RQ4 and RQ6, composite ACT score was shown to be significantly related to MoGEA subtest scores. Because the minimum required composite ACT score for college/university admission differentiates moderately selective institutions from selective institutions, the findings of RQ1 also connect to these results; teacher candidates from selective institutions, generally having higher composite ACT scores, are scoring higher on the five MoGEA subtests as well. The researcher recommends EPP personnel, especially those at moderately selective institutions, develop and implement strategies to support and prepare teacher candidates whose composite ACT scores are near the minimum score required for college/university admission.

In addition to the use of composite ACT score as a predictor of success on the MoGEA, EPP personnel should also consider the following demographic characteristics when working to identify which teacher candidates might require additional support and preparation: ethnicity, gender, total number of credit hours earned at the time the student’s highest reported score on the MoGEA was obtained, GPA at the time the student’s highest reported score on the MoGEA was obtained, and whether the candidate received financial aid. Based upon the results of the study’s statistical analyses, EPP personnel should use the following general summations of teacher candidate performance on the MoGEA to identify teacher candidates who will potentially require additional support and preparation prior to the examination:
• Teacher candidates who self-identify as Black, Non-Hispanic or Hispanic/Latino obtain lower scores on the MoGEA than teacher candidates from other ethnic groups.

• Female teacher candidates obtain lower scores on the MoGEA than male teacher candidates on every subtest except the Writing subtest. Conversely, EPP personnel should note male teacher candidates are not as successful as female candidates on the Writing subtest.

• As the total number of credit hours teacher candidates have earned increase, so do their MoGEA scores. Possibly because the MoGEA was designed to assess basic knowledge following the completion of the 42-hour general education requirement, students who have taken and completed more classes perform better on the MoGEA. EPP personnel should consider encouraging students to wait until they have completed the full 42-hour general education requirement before taking the MoGEA.

• As teacher candidates’ cumulative GPAs increase, so do their MoGEA scores. Especially if the coursework a teacher candidate has completed at the time she or he plans to take the MoGEA is predominantly composed of general education courses, EPP personnel should identify students with low GPAs as potentially requiring additional support and preparation prior to the examination.

• Teacher candidates who do not receive financial aid obtain lower scores on the MoGEA than teacher candidates who do receive financial aid. The reasons for this are unknown, however, and this finding warrants further research.
Recommendations for Future Research

The purpose of this quantitative study was to recommend statewide cut score norms for the MoGEA to MACTE for the purpose of informing its representatives to MACCE. Additionally, this study aimed to identify demographic characteristics of EPP candidates that might serve as predictors of success on the MoGEA. The following recommendations for future research are based upon the study’s data collection, review of literature, findings, and conclusions.

1. The Missouri BOE, DESE, the Commissioner of Education, and EPP personnel statewide should explore means other than basic knowledge testing to determine teacher candidate admission into Missouri’s EPPs. The conclusions of this study, as well as those of past studies related to standardized testing and teacher certification, reveal using standardized tests as qualifiers for EPP admission disproportionately eliminates Black and Hispanic teacher candidates from the teaching workforce. Especially considering the aforementioned research supporting the need for more diversity in the teaching force and the educational benefits minority students experience when taught by minority teachers, the MoGEA should not be used as a qualifier for teacher candidacy.

2. Despite numerous attempts to solicit participation from EPPs throughout the state of Missouri, only ten of the approximately 50 EPPs submitted teacher candidate MoGEA scores for the study. The absence of data from the vast majority of the state’s EPPs necessitates further study in order to more fully understand the possible effects of setting statewide MoGEA cut score norms.
Researchers and EPP administrators should consider the continued data collection and analysis of teacher candidates’ MoGEA scores for the following populations:

- teacher candidates from open enrollment or highly selective institutions;
- teacher candidates from historically black colleges/universities; and
- teacher candidates who self-identify as an ethnicity other than White, Non-Hispanic.

3. There is a lack in literature specific to the history of teacher credentialing in the state of Missouri. Researchers, Missouri’s EPP personnel, and DESE should consider creating an historical account of teacher credentialing in Missouri to provide education professionals with a greater understanding of the developments, decisions, and events that led to the current credentialing process and requirements necessary to join the ranks of Missouri’s teachers.

4. Researchers and EPP personnel should further investigate many of the current study’s findings regarding differences in various population subgroups’ performance on the MoGEA subtests. This study was designed to identify if significant differences in performance existed, but it did not aim to identify reasons for the significant differences. Further research should be conducted to determine the reasons for these differences so EPP personnel can better support and prepare teacher candidates.

5. EPP personnel should develop and implement strategies to both identify and support teacher candidates who have been shown through the statistical
analyses of this study to obtain lower scores on the MoGEA than their peers. EPP personnel should also consider restructuring the coursework of teacher candidates, specifically when they complete the majority of their general education courses, to more advantageously prepare students for success on the MoGEA.

6. Because of the direct relationship between total number of credit hours earned by teacher candidates and their MoGEA scores, researchers and EPP personnel should further explore if there is any specific coursework that leads to better performance on the MoGEA. Research should be conducted to determine if more general education coursework or coursework in a specific content area or discipline is responsible for differences in performance.

**Summary**

The MoGEA was implemented as the qualifying test of basic knowledge for admission into EPPs throughout the state beginning in September of 2013. Since that time, EPPs have set their own cut scores for the five subtests and will continue to do so until September of 2015 when DESE sets cut scores to be used statewide. The purpose of this quantitative study was to recommend statewide cut score norms for the MoGEA subtests to the executive board of MACTE for the purpose of informing its representatives to MACCE. Additionally, this study aimed to identify demographic characteristics of EPP candidates that might serve as predictors of success on the MoGEA. If predictor variables are determined, EPP administrators can identify which of their education students will most likely require additional support and preparation for the MoGEA. This information will also make DESE and Missouri’s State Board of
Education aware of which EPP candidate demographics are likely to be denied EPP admission.

The statistical findings of the study revealed significant differences in how teacher candidates from different demographic groups performed on the MoGEA subtests during the 2013-2014 academic year. All null hypotheses associated with the study (RQ2, RQ4, RQ5, and RQ6) were rejected by the researcher as test biases against specific populations were revealed and predictor variables emerged. Test biases against teacher candidates from the Black, Non-Hispanic and Hispanic/Latino populations and also against female teacher candidates were of primary concern to the researcher. Composite ACT score emerged as the strongest predictor variable of success on the MoGEA subtests, but other demographic characteristics could also be used by EPP to identify candidates who will potentially struggle on the MoGEA, including ethnicity, gender, total credit hours, cumulative GPA, and whether or not a student receives financial aid.

As was discussed in the conclusions of this chapter, the researcher ultimately recommends the MoGEA not be used to determine teacher candidate acceptance into EPPs throughout the state of Missouri because it is a biased examination and will deny EPP admission to Black, Hispanic, and female candidates at rates disproportionate to those of their peers. If the MoGEA is to serve the purpose for which it was created without unfairly preventing members of specific groups of teacher candidates from becoming state-certified teachers, DESE should ensure this qualifying examination be further reviewed, developed, and field-tested and appropriate changes be made.
References


Educator Assessments, B3Z13010 (September 19, 2012).


APPENDIX A

International Review Board Exemption Letter

Campus Institutional Review Board
University of Missouri-Columbia

483 McReynolds Hall
Columbia, MO 65211
PHONE (573) 882-9585
FAX (573) 884-0663
E-MAIL umcresearchcirb@missouri.edu
WEB: www.research.missouri.edu/cirb/

September 3, 2013

Jeffrey Edmonds
800 University Drive
Maryville, MO 64464

Dear Mr. Edmonds,

Based on the information you have provided us regarding your research with secondary data that will have been stripped of all identifiers, I have determined that you are not conducting research with human subjects because:

- The investigator is not obtaining data through intervention or interaction with living individuals.
- The investigator is not obtaining data through identifiable private information.

There are no further MU IRB requirements for this project entitled “Missouri General Education Assessment: Determining Statewide Cutoff Scores and Predictor Variables”. Please notify the IRB if your role changes in ways that could affect the determination.

If you have questions, please feel free to contact the IRB office 882-9585.

Sincerely,

Heidi Mitchell, B.A.
Compliance Specialist
VITA

Jeff Edmonds is currently in his eighth year of teaching at the Catherine Cook School in Chicago, IL, where he happily teaches middle school mathematics and coaches basketball and volleyball. Prior to teaching mathematics, Jeff directed choirs at Sacred Heart Preparatory Academy in Atherton, CA, and also served as the music director for K-12 musical theater productions. Jeff’s educational interests include technology integration, mathematics pedagogy and assessment, and adolescent psychology. Jeff is an avid reader, cook, athlete, and musician.