Teacher Perceptions of the Ceiling Effect With Gifted Students and the Impact on Teacher Value-Added Scores and Teacher Evaluation

Brian Billings

A Dissertation
Submitted to the Graduate Faculty of The University of Findlay’s College of Education in partial fulfillment of the requirements for the degree of

DOCTOR OF EDUCATION

May 2017

Committee:

Jon Brasfield, Ph.D., University of Findlay
Chair, Dissertation Committee

Nicole V. Williams, Ph.D., University of Findlay
Committee Member

Jim Fritz, Ed.D., Anthony Wayne Local Schools
Committee Member
ABSTRACT

The purpose of this study was to examine teacher perceptions of the ceiling-effect with their gifted students and how this impacted those students' ability to show growth with their value-added scores on state assessments. Furthermore, the impact that failing to show growth in terms of value-added scores had on teachers’ evaluations was explored. Since teachers of math in grades 4-6 now have their students' scores used as part of their evaluations, this study was conducted at a great time in terms of sharing the results with those in the educational community. A quantitative, non-experimental approach was taken that focuses on teacher responses on a survey that includes their perceptions, value-added scores and teacher evaluation rating. Due to concerns with confidentiality, all identifiable information was removed from these responses.

Based on this study, regular classroom teachers with high quality professional development had the highest average value-added score which was even higher than replacement courses taught by gifted intervention specialists. Teacher perceptions of working with gifted students showed no significant effect on student growth as seen in value-added scores or teachers' summative OTES evaluation rating.

Keywords: gifted students, value-added scores, teacher evaluations, gifted service models, ceiling effect, above-level testing.
DEDICATION

This dissertation is dedicated to my wife Taylor, my son Hugo, my mother Maureen, my father Frank, my sister Katie, my mother-in-law Debbie, my father-in-law Scot, my sister-in-law Jordan and especially to my late brother Matthew.
ACKNOWLEDGMENTS

I would like to show my gratitude and appreciate for the hard work that my dissertation chair, Dr. Jon Brasfield, participated in throughout this process. His guidance was instrumental in initially designing and ultimately completing this research. I hope he enjoyed the process and finds the final copy to be an authentic product of the work we participated in over the past three years.

In addition, I would like to thank the members of my dissertation committee, Dr. Nicole Williams and Dr. Jim Fritz. Their insight and encouragement helped me push on and complete this research. They both offered distinct perspectives that helped enrich the final product and my time as a doctoral student.

The University of Findlay has provided an excellent doctoral experience that I will always be proud of and promote to others. The work of the program chairs, Dr. Michael Scoles and Dr. John Gillham, set the tone for the positive and challenging experience this program would prove to be for me both personally and professionally. All of the professors brought a unique perspective and student-first mentality to their courses. The various support staff was always positive and helpful when approached. In addition, my two years as a research assistant was beneficial to both my writing and financially in paying more this degree.

The success of any person depends mightily on those who support them in life. I have had the tremendous honor of having a great group of professionals to work with and mentor me. This began with my tutelage under Anne Heinl at Sylvania Arbor Hills Junior High during both my methods and student teaching. I am so thankful she took a chance on another student teacher and how demanding she was of the work we completed. I will forever be grateful to Les Schultz, Brenda Gift and the whole ESC of Lake Erie West organization for offering me my first job and helping to develop my teaching ability. My time working under Serena Troyan helped me
transition from a green, new teacher to a passionate, productive educator. Her advice and guidance is still in effect today when I make decisions.

I have had the honor and privilege of working with two tremendous public school systems. As an Anthony Wayne graduate, the chance to work in the district and impact the lives of hundreds of students as been both inspiring and humbling. There is no greater profession in my mind than teaching and altering the path of youth. The leadership of Dr. Jim Fritz has infused this passion into all the staff in the district. The daily support offered by Kevin Herman along with the other staff and administrators truly embodies the notion of a district team. The opportunity to partner with Perrysburg Schools in overseeing gifted education was an unexpected but rewarding adventure. Tom Hosler is a thoughtful leader with a sense of humor that can lighten any room. His willingness to bring me into the district and value my opinion settled me into my position rapidly. There is perhaps no more passionate and skilled education professional than the future Dr. Kadee Anstadt. She is an inspiration to me on a daily basis and someone I hope embody in my career. The Teaching and Learning Department is in good hands with her and the excellent staff.

As well as my professional supports, I am thankful for the amazing people in my life who continue to challenge and shape me on a daily basis. Lori Williams continues to set the standard for what it means to be exceptional in everything someone does. The sheer number of people and organizations that have impacted me would be too great to name. The support offered to me by my parents allowed me the opportunity to have a great education and set me up for success in life. I will be forever grateful for their sacrifice for my benefit. The addition of my in-laws was both an unexpected and amazing change. Their acceptance of a weird 18 year old to their family changed the trajectory I was on and led me down the path I am on today.
Finally, I would like to thank my wife, Taylor, for her patience and encouragement in pursuing my doctorate. A dissertation is truly a burden on both the writer and their spouse as time together is sacrificed to complete the work. Her motivation helped me work through the tough legs of this marathon and finish my research. Her assistance is even more impressive as we added our son, Hugo, to our family in the middle of my work. The short term sacrifices of missing time with him will be worth it now that I have finished. He has been a tremendous blessing on our family and we can not imagine life without him.
TABLE OF CONTENTS

Page

CHAPTER I: INTRODUCTION ............................................................................................................... 1
  Background of the Problem........................................................................................................ 3
  Statement of the Problem.......................................................................................................... 8
  Rationale of the Study.............................................................................................................. 8
  Purpose of the Study................................................................................................................ 9
  Significance of the Study........................................................................................................ 11
  Research Questions................................................................................................................ 12
  Definition of Terms................................................................................................................ 13
  Limitations.................................................................................................................................. 14
  Delimitations............................................................................................................................. 14
  Researcher Bias ...................................................................................................................... 15
  Conclusions............................................................................................................................... 16

CHAPTER II. LITERATURE REVIEW .............................................................................................. 17
  Gifted Students and Scoring (Ceiling Effect)........................................................................... 20
    Use of Above-level Testing for Gifted Students..................................................................... 21
  Teachers and Their Impact on Student Achievement.............................................................. 23
  Implications of Using Value-added for Teacher Evaluations................................................... 24
  Teacher Evaluation and Models.............................................................................................. 25
  Teacher Perceptions of Gifted Students.................................................................................... 26
  Gifted Service Models............................................................................................................. 27
  Summary..................................................................................................................................... 28

CHAPTER III. METHODOLOGY ..................................................................................................... 30
Research Questions........................................................................................................ 30
Research Design........................................................................................................... 31
Participants................................................................................................................... 31
Instrumentation & Data Sources.................................................................................. 33
Data Collection Procedures........................................................................................ 35
Data Analysis................................................................................................................ 36
Assumptions................................................................................................................... 39

CHAPTER IV. RESULTS.............................................................................................. 42
Characteristics of the Sample....................................................................................... 42
Instrument Validity and Reliability............................................................................. 43
Research Question 1..................................................................................................... 44
Research Question 2..................................................................................................... 46
Research Question 3..................................................................................................... 48
Research Question 4..................................................................................................... 49
Research Question 5..................................................................................................... 51
Summary....................................................................................................................... 53

CHAPTER V. CONCLUSIONS AND RECOMMENDATIONS................................. 55
Review of the Study....................................................................................................... 55
Discussion...................................................................................................................... 57
    Research Question 1................................................................................................. 57
    Research Question 2................................................................................................. 58
    Research Question 3................................................................................................. 59
    Research Question 4................................................................................................. 60
    Research Question 5................................................................................................. 60
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conclusion</td>
<td>61</td>
</tr>
<tr>
<td>Recommendations</td>
<td>63</td>
</tr>
<tr>
<td>Future Research Opportunities</td>
<td>64</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>66</td>
</tr>
<tr>
<td>APPENDIX A. IRB APPROVAL</td>
<td>69</td>
</tr>
<tr>
<td>APPENDIX B. INVITATION TO PARTICIPATE IN THE STUDY</td>
<td>70</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>OTES Original Framework for Teacher Evaluation</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>OTES Alternate Framework for Teacher Evaluation</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>OTES Final Summative Evaluation Rating for Teacher Evaluation</td>
<td>4</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Graduation Points Earned by Performance Level on End of Course Tests</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>Math Achievement Scores for Ohio Students for 2006-07</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>Research Question Summary</td>
<td>39</td>
</tr>
<tr>
<td>4</td>
<td>Number of Respondents, Mean Value-Added Score and Variance by Service Model Type</td>
<td>45</td>
</tr>
<tr>
<td>5</td>
<td>Number of Respondents Concerned or Very Concerns About the Ceiling Effect and the Methods Used to Mitigate the Impact</td>
<td>47</td>
</tr>
<tr>
<td>6</td>
<td>Number of Respondents, Mean Value-Added Score and Variance Based on Teacher's Perception</td>
<td>49</td>
</tr>
<tr>
<td>7</td>
<td>Number of Respondents, Mean Summative OTES Evaluation and Variance Based on Teacher's Perception</td>
<td>51</td>
</tr>
<tr>
<td>8</td>
<td>Number of Respondents, Mean Value-Added Score and Variance Based on Teachers' Summative OTES Evaluation Rating</td>
<td>52</td>
</tr>
</tbody>
</table>
CHAPTER I. INTRODUCTION

Introduction

Since 2008, there has been a dramatic rise in the examination of teacher value-added scores and how it relates to teacher performance both in Ohio and around the country. This has put a renewed focus on students’ scores on standardized tests and how teachers can help students show progress from year to year. While this has added a new level of stress for teachers, this has been a welcome shift from past focuses on proficiency, at least in terms of the needs of gifted (high-ability) students. Having been put through the No Child Left Behind (No Child Left Behind, 2001) era and its focus on meeting proficiency with all students, Ohio has now transitioned to a focus on showing growth (as determined in value-added scores) for each student. Regardless of the level that any student starts at, every student must be shown to have had a productive year of instruction that has helped them grow. Every student is now examined independently rather than just having two groups of students: those who met proficiency and those who did not. In terms of gifted students, some might see that as a positive move as no longer can the top performing students be neglected in an effort to focus on those struggling for proficiency. The effect of this change though has been that teachers now worry about not just meeting the needs of all students in their class appropriately and effectively as they see fit, but actually having data to support these claims.

The era of NCLB brought national attention to the issue of standardized testing and gaining proficiency for students. Teachers were in the midst of these debates and were tasked with bringing the bottom up in terms of classroom performance. Research by Loveless, Parkas, & Duffett (2008) supports the idea that teachers believe that focusing on bringing the standardized test scores of underperforming students to proficiency has pulled attention and resources away from higher-achieving students. Few teachers have positive things to say about
the impact of NCLB on academically advanced students (Loveless, Parkas, & Duffett 2008). In
fact, when asked to either agree or disagree with the statement that “getting underachieving
students to reach ‘proficiency’ has become so important that the needs of advanced students take
a back seat”, more than three in four (77%) agreed (Loveless, Parkas, & Duffett 2008). The
work by Loveless, et al. shows that this focus on proficiency is a very real concern among
educators and parents of high-ability students. If the focus is simply on achieving proficiency,
the needs of the high-ability students will be neglected. High-ability students have unique needs
that need to be challenged in order for them to show academic growth. A focus on the lowest
performing students and achieving proficiency is a major hurdle for teachers with gifted students.
In addition, how to identify a gifted student based on their visible traits and what their
background is can be a further challenge when discussing these tests.

To help clarify how gifted students are identified, the researcher will use the National
Association for Gifted Children’s (NAGC, 2010) definition of giftedness that sets the standard as
follows:

Gifted individuals are those who demonstrate outstanding levels of aptitude (defined as
an exceptional ability to reason and learn) or competence (documented performance or
achievement in top 10% or rarer) in one or more domains. Domains include any
structured area of activity with its own symbol system (e.g., mathematics, music,
language) and/or set of sensorimotor skills (e.g., painting, dance, sports) (p. 1).

Furthermore, gifted students tend to have a distinct family and socioeconomic
background when compared to the population as a whole as determined by information collected
on various assessments. The National Assessment of Academic Progress (NAEP) is the largest
assessment given to a nationally representative sample of American students of what students
can do in multiple subjects including mathematics and reading (National Center for Education
The typical student scoring at the 90th percentile on the eighth-grade math NAEP comes from a more privileged socioeconomic background than the typical American student. Only 10.2% qualify for free or reduced price meals, compared to 36.1% of eighth-graders nationwide and 66.5% of students scoring at the 10th percentile (National Center for Education Statistics, 2011). This means that high achievers are only one-sixth as likely to be eligible for the free or reduced price meals program—a proxy for family income—as low achievers (Loveless et al., 2008). In addition, students who receive free or reduced lunch were less likely to even be recommended for testing by their teachers when compared with those students who pay for their own (McBee, 2006). This tends to indicate that students from a wealthier background are more likely to even be tested than their poorer counterparts.

**Background of the Problem**

For Ohio reading and/or math teachers in grades 4-8, the need to meet growth with their high-ability students’ test scores on state assessments is part of the job. Recently, changes in the Ohio Legislature have created the Ohio Teacher Evaluation System (OTES) rating system that places teachers in one of four categories: Accomplished, Skilled, Developing or Ineffective. For the teachers listed above, 50% of their evaluation is based on their students’ test scores on the state assessments though recent legislation has given districts the option of reducing this percent to 42.5% (Ohio Department of Education, 2016). Figure 1 outlines the original framework set-up, Figure 2 presents the alternative framework set-up, and Figure 3 presents how the final summative rating is calculated.
Figure 1. OTES Original Framework for Teacher Evaluation

Figure 2. OTES Alternate Framework for Teacher Evaluation

Figure 3. OTES Final Summative Evaluation Rating for Teacher Evaluation

The evaluation factors are weighted as follows:

1. If a district chooses the original framework, the teacher performance measure and student growth measure shall be 50% each.
2. If a district chooses the alternative framework:
   - The teacher performance measure shall account for 50%;
   - The student academic growth measure shall account for 35%; and
   - The chosen alternative component(s) shall account for 15%.

Figure 4. OTES Final Summative Evaluation Rating for Teacher Evaluation Tables
Every teacher now needs to show growth with the average from their entire class of students which includes the top performer down to the struggling one. This can create a very intense classroom atmosphere that is scaring some teachers away from working with gifted students for fear that they will be unable to help these students show growth. This perception among teachers is something that requires further investigation. Their potential fear is a phenomenon referred to as the “ceiling effect” that can be common with assessments for gifted students.

For gifted students, an on-grade level assessment may not be able to truly differentiate between their levels of mastery. This problem with a test’s threshold is referred to as the ceiling effect. According to Koedel & Betts (2010), the ceiling effect can be seen as the tendency for growth in a student’s test score to be less if their initial score is at the top end of the scores as the student has little room for growth given the difficulty level of the test. The ceiling effect is defined as follows:

Operationally defined as the clustering of scores at the upper limit of the test. Group aptitude and achievements tests and some individual tests are simply too easy for highly gifted students. Ninety-ninth percentile scores on annual tests may please educators and parents, but they do not yield a true picture of the specific functioning of the child. For the highly gifted child, grade-level test scores tell only the percentage of students that performs below the individual but obscure what the child could have achieved had the test included appropriately difficult items. The problem is analogous to that of trying to measure the heights of 12 year old children using a measuring stick that is only 5 feet long. Many children can be measured using the stick, but we cannot differentiate among those who are 5 feet tall and those who are almost 7 feet tall (Stanley, 1990, p.166).

As a practical example of the ceiling effect, let’s say two students take a 5th grade math assessment. Student A may be at the 99th percentile and have gotten every question right (and
possibly be able to answer questions above their grade level). Student B might have missed a
couple questions and just barely passed the threshold for the same percentile. These two students
are remarkably different yet the test would inform a teacher they are achieving at nearly the same
level. Without being able to truly assess both students true level, the teacher might struggle to
provide the targeted instruction both students need. This leads educators to think the test scores
of gifted students are accurate measures of their abilities because they have no opportunity to
observe the true capabilities of these students. This challenge is outlined by Koedel & Betts
(2010) when discussing the fact that ceiling effects will be most severe in minimum competency
or proficiency-based tests. This was especially true during the No Child Left Behind (NCLB) era
as the focus was largely on proficiency.

As noted by Silverman (2012) in an article on the assessment of gifted students, test
ceilings have a tremendous amount of difference with grade-level group achievement tests
having low ceilings as compared to the higher ceilings on individual IQ tests. Warren, J.R.,
Grodsky, E. & Lee, J.C. (2008) make this point in their article discussing the use of high school
exit exams. As they point out, 22 different states nationwide use high school exit exams that are
created at a middle-school or lower high-school level rather than a more rigorous level. While
Ohio was not mentioned specifically in this article, end of course exams have now taken hold in
Ohio even impacting graduation requirements. As seen in Table 1, students now have to earn 18
points from a possible 35 points on the seven end of course exams in Ohio (Ohio Department of
Education, 2016).
Table 1. *Graduation Points Earned by Performance Level on End of Course Tests*

<table>
<thead>
<tr>
<th>Performance Level</th>
<th>Graduation Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced</td>
<td>5</td>
</tr>
<tr>
<td>Accelerated</td>
<td>4</td>
</tr>
<tr>
<td>Proficient</td>
<td>3</td>
</tr>
<tr>
<td>Basic</td>
<td>2</td>
</tr>
<tr>
<td>Limited</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: Ohio Department of Education

In these instances of deciding whether the ceiling effect is in play, a 4th grade math teacher might be worried when they get a class roster that includes standardized test scores for a student who scored in the 99 percentile the previous year. The math teacher will likely think about the struggle to show progress again this year and how to help this student who clearly is gifted in math. This can lead to frustration and a feeling of hopelessness for the teacher who may not have the resources and support systems in place at school to work with these students. Knowing that their evaluation is tied to this student's score could be the cause for much frustration. In this case, a teacher's perception of working with a high-ability student could be possibly problematic and negatively impact student performance.

The current literature is lacking for the ceiling effect in terms of value-added scores and the impact on teacher performance. There has been plenty of focus on topics such as the ceiling effect (Stanley, 1990), concepts of giftedness and common traits of gifted students (Loveless et al., 2008). While the ceiling effect is a recognized problem (Stanley, 1990) with assessments for high-ability students, there is a lack of available research on the implications of gifted students suffering from the ceiling effect in terms of how this affects teacher value-added scores. The issue remains to be seen if this is an actual phenomenon and if so, what are the impacts on
teachers of these students. Is it fair to expect teachers to show the same amount of growth with the top 10% of students as the bottom 10%? If not, what else can be done to gather a more complete picture of these students’ true abilities? There is research indicating that value-added estimates are generally immune to ceiling effects (Koedel & Betts, 2010) especially when using norm-referenced tests rather than a proficiency test.

To counter the ceiling effect, a fair amount of research has been done on using above grade level tests with high-ability students. Hansen (1992) provides guidelines to using above-level testing including providing recommended tests to use and what type of data should be selected to provide an accurate picture of the student’s abilities. Others, such as Van Tassel-Baska (1986), encouraged using the commonly given SAT as an above-level test and diagnostic tool.

Statement of the Problem

This study addresses the issue with teacher perceptions of working with gifted students in math in grades 4-8 topping out, or incurring the ceiling effect, on grade-level achievement tests and the impact on teacher evaluation ratings. The ceiling effect clouds the true picture of how they are performing and limits the knowledge the teacher gains from their scores. Teacher perceptions of gifted students and whether they believe they can even show growth will be examined to see what, if any, impact that has on student scores. The impact not showing growth with these students has on their teachers’ value-added scores and ultimately, their teacher evaluation rating will be examined.

Rationale

Today’s educational culture in Ohio relies heavily on data to determine the quality of the education that is happening. This data is now being partially used to determine the level of performance among teachers. Since growth needs to be shown for all students, this can bring
apprehension to a teacher who has many gifted students assigned to their class. If a teacher has a class full of students who are already scoring at the top of their state tests, they may feel like it is not possible to show growth with their students. Previously, the target was simply proficiency which meant that high ability students had to learn very little in order to help a district reach their goals (Braun, 2005). Now however, value-added models have taken hold so all students must show growth.

People could point to specialized programs such as pull-out programs and accelerations (subject and grade) available for high-ability students as examples of how the needs of these students can be met. While that may have some merit, these classes can often be hit and miss in terms of meeting the needs of gifted students. As examined by Loveless et al. (2008), “in fact, 40 percent of teachers say that the content and curriculum of honors and accelerated classes is too often watered down and lacking rigor”. Regardless, the struggle with teacher perceptions of the ceiling effect negatively impacting their high-ability students could mean that they may see the challenge to show growth as insurmountable.

Since gifted students can suffer from the ceiling effect if taking a proficiency based test, or the teacher believes that their gifted students will struggle on any type of test, schools must take strides to make sure these students are put into the right classroom setting both in terms of regular education classes and any special courses like pull-out programs. Teachers’ ratings are now partially dependent on the growth of these students. To make sure these ratings are as fair and accurate as possible, all students should be placed in the right educational setting that gives them, and by extension their teachers, the maximum chance to grow.

**Purpose of Study**

The purpose of this study is to examine whether a teacher's perceptions of showing growth with gifted students as defined by value-added scores has an impact on the students'
actual scores. In addition, the study focus on the service models used by districts and if they have an impact on student scores. This will also lead to an examination into whether a lack of student growth for high-ability students is negatively impacting their teachers' value-added scores and teacher evaluation. The move to value-added models has been a positive trend in terms of providing opportunities to discuss education. As noted by Braun (2005), “value-added modules move the discussion about teacher quality to where it belongs: centered on increasing student learning as the primary goal of teaching” (p. 6).

In terms of teachers and administrators involved in gifted education, one of the key components of helping a high ability student grow is ensuring they are in the proper environment. Ensuring gifted students are challenged is crucial so recognizing when a student is topped out in their current environment (as demonstrated by the test scores) becomes a challenge. This study will provide evidence of the ceiling effect and how those involved in gifted education can be best serving their students.

In addition, another reason to be excited by this study is that it found a unique niche in the current research. There has been a fair amount of work done regarding intelligence tests, the ceiling effect and the growth of high ability learners; however, there is a gap in the research when it comes to teachers' perceptions of working with gifted students and how this affects teachers’ value-added scores in today’s high stakes environment. There is the potential for 4th to 8th grade reading and math teachers to be very nervous about having gifted students in their class. They could be worried about how they can help these students show growth if they are already scoring so high. This dilemma is certainly understandable and this study will be a great resource for these discussions.

In terms of stakeholders who might be interested, there would be a large audience of interested parties. District leaders and building principals who are dealing with the gifted
indicator and other progress measures for these students would be interested along with classroom teachers whose value-added will be tied to these students. In terms of researchers, Dr. Joyce VanTassel-Baska is considered the number one authority on all things gifted. The Center for Talent Development (CTD) at Northwestern University is another authority on gifted issues that does a lot of work with high-ability students and above-level testing using the ACT and SAT.

**Significance of Study**

This study takes into account the thousands of teachers in Ohio that are currently receiving value-added scores and are having their effectiveness as educators tied to this data. For the 2013-14 school year, this includes 2,179 5th grade math teachers and 2,502 5th grade reading teachers (SAS Institute, 2015). Countless hours are spent each year going through data with these teachers to help them understand how their students are scoring and what they can do to help them show progress. Furthermore, researchers are doing their best to ensure that the progress data that is published annually is fair and shows a proper reflection of the growth that is occurring. Addressing an issue that can fundamentally impact the growth being shown by students is of the utmost importance.

The study will have a meaningful impact on both those involved in gifted and talented education along with the regular classroom teachers in grades 4-8 who deal with value-added scores. In terms of gifted intervention specialists and administrators, this research can help them monitor their students’ progress to ensure they have their students in the proper settings. Even if just 1% of their population needs to have their setting adjusted, this could lead to tremendous growth for that student which will translate to the teacher. The goal with education should be to reach all students where they are and to help them raise their level of learning. This study will matter to the gifted students in Ohio whose academic and social/emotional needs should be met in their districts. My hope is to find answers in the research that can help guide instructional
practices in classrooms with teachers who are responsible for showing growth with gifted students.

For the regular classroom teacher who receives a value-added report each year, this study can serve as a self-check for them when examining their high performing students. Is one of their students showing growth and scoring well? Or has that student topped out and is being held back in their current setting? Good educators are very reflective and constantly think through what they are doing and why it is happening.

As someone who spends every day working with gifted students, testing, educational placements and programming, this study is extremely significant as I believe all too often that high ability students are not given the same focus as other students. I am tasked with collecting data and monitoring student growth measures constantly. The data-saturated education culture means that you have to be aware and understand the information at hand. Being tasked with recommending changes to improve the educational experience for students and the ability to show growth with these students by teachers is an extremely challenging position.

**Research Questions**

The impact of teachers’ value-added scores is only now being examined in terms of student growth and teacher evaluations. No study has examined the issue of value-added scores for gifted students. This study will examine the issue with gifted students in grades 4-6 in math topping out, or incurring the ceiling effect, and the negative impact it has on their teachers’ value-added scores. Specifically, the research questions that will guide the study include:

1. How is a district's service model related to teachers' progress with their gifted students in terms of value-added math scores?
2. Are districts concerned with the ceiling effect (students topping out by hitting a 99% on a test) for their gifted students, and if so, how are they attempting to minimize its impact?

3. How are teacher perceptions of having gifted students in their class related to their value-added math scores?

4. How are teacher perceptions of having gifted students in their class related to their OTES teacher evaluation ratings?

5. How do gifted teachers' value-added math scores compare to their OTES teacher rating?

**Definition of Terms**

To assist in the discussion of the study, the following key terms were defined based on a study of the relevant literature.

*Gifted/High-Ability Student*: the National Association for Gifted Children’s (NAGC, 2010) definition of giftedness sets the standard as follows:

Gifted individuals are those who demonstrate outstanding levels of aptitude (defined as an exceptional ability to reason and learn) or competence (documented performance or achievement in top 10% or rarer) in one or more domains. Domains include any structured area of activity with its own symbol system (e.g., mathematics, music, language) and/or set of sensorimotor skills (e.g., painting, dance, sports) (p. 1).

*Ceiling-Effect*: operationally defined as the clustering of scores at the upper limit of the test. Group aptitude and achievements tests and some individual tests are simply too easy for highly gifted students (Hansen, 1992).
Education Value-Added Assessment System (EVAAS): a system created by the SAS Institute out of North Carolina that creates models for tracking, assessing and predicting student performance (SAS Institute).

Value-Added Scores:

“also known as growth measures, are used to estimate or quantify how much of a positive (or negative) effect individual teachers have on student learning during the course of a given school year. To produce the estimates, value-added measures typically use sophisticated statistical algorithms and standardized-test results, combined with other information about students, to determine a “value-added score” for a teacher (Glossary of Education Reform).”

Service Model: the method used for providing gifted services to students such as pulling them out into a resource room or using acceleration.

Limitations

This research will have limitations based on the fact the state tests switched to a new vendor this past year and there are a variety of concerns with the results. Some districts used paper/pencil versions of the test while others completed online assessments. There were also many students in the state who opted out of testing this year which removed some scores from teachers data. When collecting the teacher's value-added score in the survey, these issues may pose a problem in the data. Since the research methodology will be focused on inferential statistics, there is no way to be completely sure that the values calculated are correct since the population was not fully measured. In addition, the survey has been created by the researcher so there is the possibility the instrument might be not as reliable as hoped.

Delimitations
Since gifted students are a point of focus, only teachers who have identified gifted students in their class will be included in the research. Second, studying the ceiling effect is an essential part of this study so only math teachers are included to collect information on districts who are math accelerating students. Finally, value-added is a key component of the research so only teachers in grades 4-8 will be respondents.

**Researcher Bias**

My expectations of this study are that high-ability students are not showing growth on their state assessments due not to the ceiling effect of the test, but rather the perceptions of the teachers. Teachers that are providing daily instruction in a specific subject will be showing more growth than those who are simply offering enrichment. The assumption is that districts that are using accelerations (moving students up a whole grade level or just in one subject) for gifted students will be showing more growth than those keeping these students in their original grade. The claim that accelerations are the best way for districts to minimize the ceiling effect for their high-ability students is being studied.

With the new focus on student scores in terms of their teacher’s evaluations, the impact of gifted students showing growth is essential for those teachers wishing to hit the accomplished (or even skilled) level. High-ability students will be a huge challenge for teachers as they look towards their student scores. In a class with a small cluster of gifted students, the impact will be manageable as teachers have other students that have the potential for tremendous growth. A handful of students at the ceiling effect will be lessened by the remainder of the class assuming the teacher can help those students show growth. In a class with a large percentage of the students being identified as gifted, there will be a large impact on the teacher’s scores as there are fewer regular students to offset the ceiling effect.
In terms of bias, my daily experience working with gifted students provides me much information about these students. This information also leads me to have some potential research bias in my study. I am assuming there is a problem with student’s scores and that they are hitting the ceiling effect. Since there is so little research regarding teacher’s value-added scores in terms of their gifted students, there are many unanswered questions to explore. In the end, the data will do the talking.

Conclusion

Recent changes in Ohio have brought gifted students to the forefront of the educational climate. In particular, gifted students’ value-added scores have become a point of focus in terms of how districts and teachers are performing. This study has set out to explore the impact of the ceiling effect and teacher perceptions on gifted students’ value-added scores and teacher evaluation. Furthermore, does the ceiling effect, or teachers perceptions of the ceiling effect, impact teachers’ value-added reports as their gifted students show minimal growth. Different service models and how they impact students’ scores were explored to give a picture of what works best in order for high-ability students to show growth. What type of service model was used and the corresponding growth scores provided an interesting look at what services work with high-ability students.
CHAPTER II. LITERATURE REVIEW

Introduction

The recent changes in the teacher evaluation process in Ohio have now made classroom teachers' value-added scores even more apparent for the general public. Previously, teachers knew that their students' scores would be viewed by administrators as one example of their instruction. Now, up to half their evaluation could be based on their students' performance on the state tests. With some teachers' unions opting for all teachers to get a shared portion of value-added scores, teachers with gifted students in their class now could impact the entire staff. This is especially true due to the fact that less than 5% of teachers feel the highest performing students have been a focus previously in terms of tracking their scores and trying to raise test scores (Loveless, et al., 2008).

One of the main concerns among educators seems to be a fear that since gifted students have previously scored highly on the standardized tests, it will be impossible for them to show growth again for the coming year. The rationale seems to be that it would be better to take an average scoring student so there is plenty of room for growth as measured by value-added. Much of this concern seems to be due to misunderstandings of how value-added is actually calculated and what type of stretch or extension the standardized tests offer. According to Koedel & Betts (2010), it is essential to differentiate between “criterion-referenced” and “norm-referenced” testing. In criterion-referenced tests, tests are set to show proficiency on a set of standards which are typically state approved. In these circumstances, it is possible for student growth with high achieving students to be lessened.

This is particularly true in situations where the standards the test is based on are not rigorous. Using Ohio as an example, for the 2006-07 school year, 86 percent of fourth-grade students scored at or above the “proficient” level in math on the state-level Ohio Achievement
Assessment (OAA). However, just 46 percent of these students scored at or above the proficient level on the National Assessment of Education Progress (NAEP). These statistics are especially sobering for minority and economically disadvantaged students. Whereas 50 percent of Black and 62 percent of economically disadvantaged students met the state's proficiency level, only 18 percent of Black and 23 percent of economically disadvantaged students met the proficiency level on NAEP (U.S. Department of Education, 2008). This trend is similar with the data for 8th graders as seen in Table 2. There is a stark contrast between Ohio's definition of proficiency and the more rigorous national assessments.

Table 2. Math Achievement Scores for Ohio Students for 2006-07

<table>
<thead>
<tr>
<th>Category</th>
<th>State Data- % Proficient</th>
<th>NAEP Data- % Basic</th>
<th>NAEP Data- % Proficient</th>
<th>State Data- % Proficient</th>
<th>NAEP Data- % Basic</th>
<th>NAEP Data- % Proficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>76%</td>
<td>87%</td>
<td>46%</td>
<td>72%</td>
<td>76%</td>
<td>35%</td>
</tr>
<tr>
<td>White</td>
<td>82%</td>
<td>93%</td>
<td>53%</td>
<td>78%</td>
<td>83%</td>
<td>42%</td>
</tr>
<tr>
<td>Black</td>
<td>50%</td>
<td>67%</td>
<td>18%</td>
<td>45%</td>
<td>47%</td>
<td>9%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>61%</td>
<td>76%</td>
<td>25%</td>
<td>56%</td>
<td>63%</td>
<td>25%</td>
</tr>
<tr>
<td>Low Income</td>
<td>62%</td>
<td>75%</td>
<td>23%</td>
<td>54%</td>
<td>60%</td>
<td>16%</td>
</tr>
</tbody>
</table>

Source: Ohio Testing Data and 2007 National Assessment of Educational Progress (NAEP) Data

In contrast, as outlined to Koedel & Betts (2010), a norm-referenced test is designed to show student performance by ranking students compared to a reference group which is often a national student population. Assuming the test is well designed, it should minimize any concerns with high-ability students not being able to show growth. This is due to the test having to include a wide range of test items to differentiate between such a large sample of students. Due to these major differences in the type of test, it is essential to be specific in the type of test being used when discussing concerns over the possibility of student growth for high-ability students.
Furthermore, many teachers do not feel high ability students are a priority at all. In a 2008 study by Loveless, Parkas, & Duffett that collected survey findings from 900 randomly selected, nationally representative public school teachers in grade 3-12, less than one in four teachers (only 23%) said the needs of high ability students are a top priority for their school with a further 32% indicating they were actually a low priority. Much of this was based on NCLB with it's focus on showing proficiency. With the recent focus moving toward value-added scores, one would expect to see a shift in attitudes towards helping all students as each students' scores now impact their teacher's evaluation.

**Literature Search Strategies**

The literature search was conducted using the University of Findlay’s Shafer Library via remote access. The “OneSearch” feature of the EBSCO database was accessed to search for peer reviewed journal articles. In addition, Google Scholar's search feature was used to search for peer-reviewed sources.

The key search words and phrases used in the EBSCO database included “gifted”, “value-added” and “teacher evaluation.” The additional key words “IQ” and “OTES” were added in the optional search category. Articles were selected based on their relation to the topic at hand and relevancy for this discussion.

The types of literature chosen for this review include both quantitative studies and qualitative studies. Most of the quantitative studies refer to test scores on recognized standardized tests and teacher survey results, while the qualitative studies include case studies of talent search programs and teacher perceptions of high-achieving students in their schools.

**Research Questions**

1. How is a district's service model related to teachers' progress with their gifted students in terms of value-added math scores?
2. Are districts concerned with the ceiling effect (students topping out by hitting a 99% on a test) for their gifted students, and if so, how are they attempting to minimize its impact?

3. How are teacher perceptions of having gifted students in their class related to their value-added math scores?

4. How are teacher perceptions of having gifted students in their class related to their OTES teacher evaluation ratings?

5. How do gifted teachers' value-added math scores compare to their OTES teacher rating?

**Gifted Students and Scoring (Ceiling Effect)**

One of the often cited concerns with being tied to gifted students' scores is the concern from teachers that they could be subject to what is known as the ceiling effect. The ceiling effect is the occurrence of high scoring students clustering at the upper limit of any test. It can mask a true student ability assuming they had been provided more challenging items that would have provided a clearer picture of their ability (Hansen, 1992). According to Silverman (2012), in order to deal with the possible effects of the ceiling effect, it is recommended that high ability students are given tests with enough stretch to allow them to show their true ability. The ceiling effect increases with age as older students are more likely to outstrip the capacity of a test. Ceilings vary by test with group achievement tests having low ceilings and individual IQ tests having higher ones. According to Koedel & Betts (2010), norm-referenced tests are less susceptible to the ceiling effect than criterion-referenced tests such as state achievement tests.

The era of No Child Left Behind put a focus firmly on helping students achieve proficiency. While helping struggling students succeed is a noble venture, attaining proficiency is not a beneficial strategy for high-achieving students. In a 2008 study focusing on the results
from the National Assessment of Educational Progress (NAEP) testing, the gains of students scoring at the 10th percentile was 1.55 at the end of NCLB compared to just 0.35 prior to the NCLB era. This means that students who had been scoring at the 10th percentile, or better than 10 percent of other students, saw a gain in their score to the 11.55 percentile (scoring better than 11.55 percent of students). In contrast though, the gains of students scoring at the 90th percentile were just 0.625 at the end of NCLB compared to 0.675 prior to the NCLB era (Loveless, et al., 2008). Thus, while low achieving students made sizable gains, gifted students' scores plateaued as measured by the NAEP assessments.

**Use of Above-Level Testing for Gifted Students**

In response to the concerns of the ceiling effect, researchers in the field of gifted education advocate the use of above-level testing for gifted students. As discussed by Stanley (1990), if using a grade-level test for high-achieving students is analogous to trying to measure two children taller than five feet with a five foot ruler, the solution is to use a longer stick. In this case, the solution is to use a test normed for students older and/or in higher grades than the students in question. Highly gifted students can be administered above-level tests to get a more accurate picture of their true achievement. In a study conducted by Van Tassel-Baska (1986), students who scored at the 99% on a grade-level achievement test were then given the Scholastic Aptitude Test (SAT). On the math test, 21% of these students scored in the 310-400 range while 6% scored in the 610-700 range. For the verbal test, 39% of these students scored in the 310-400 range while 2% scored in the 610-700 range. At that time, the SAT score ranges went from 200-800 in both reading and math (400-1600 total). This above-level testing helps distribute students who scored at the same level in their grade to provide a more accurate view of their true ability.

This approach of above-level testing has been adopted by talent search programs such as
Northwestern University's Midwest Academic Talent Search (NUMATS). According to the Northwestern University: Center for Talent Development's (2015) website, “NUMATS is a research-validated program that utilizes above-grade-level assessment, as a means of gifted testing, to help parents and educators better understand their students’ educational needs.” NUMATS is one of many talent search programs that highly gifted students can pursue if they wish to get a better understanding of their true abilities. In many cases, the SAT is the test of choice for high performing junior high students. By giving them such a difficult test with a high ceiling, students' scores will be spread out far more than on a traditional test (Silverman, 2012).

When using above-level tests, certain guidelines have been established to ensure the most reliable and valid results. Recommendations made by Hansen (1992) as established in best practices indicates ensuring tests have sufficient ceilings to challenge students or using above-level testing as needed to ensure an accurate picture of a student's true ability. The SAT and the Stanford-Binet Intelligence Scale are two suggested tests for these situations. Regardless of the test, the scores should be used to plan program options for students. Program options include honors classes, enrichment seminars, classes at universities, grade acceleration, AP classes, and more (VanTassel-Baska, 1984).

A new twist on above-level testing has been the introduction of computer adaptive testing (CAT) where students are given initial items and then the system provides them tougher questions as they progress (or vice versa). As opposed to linear or fixed form testing where students take the same questions, CAT testing can provide a testing experience for students that is more individualized to highlight their current level of performance. Studies such as that done by Davey (2011) explore the various types of adaptive testing approaches including random form where questions are pulled randomly from a large sample, multi-stage where a small group of questions are given and the results then place the student in a new setting or even item adaptive
where each result impacts the next question (Davey, 2011). More detailed research studies such as that done by Gershon (2005) have looked in-depth at the various types of CAT and how tests are created and by Colwell (2013) which explored the implications for adaptive testing in the age of the Common Core standards.

**Teachers and Their Impact on Student Achievement**

As outlined in studies such as the ones done by Wright, Horn & Sanders (1997), Walsh (2001) and Hanushek (1971), the single most important factor affecting student achievement is teacher effect. With this in mind, it is important to understand the impact teachers have on student achievement. According to research by Sanders and Rivers (1996), teacher effect on student achievement is additive and cumulative. Their study examined math achievement scores from students over a three year span (grades 3 to 5) in two of Tennessee's larger metropolitan school districts. By dividing teachers into 5 groups (least to most effective), they could examine the impact teachers had on student performance both in a year and compounded over time. This shows that having better teachers, and having them more often, will be a bonus in terms of students' scores (some up to 50 percentile points higher). In addition, the top performing teachers help show appropriate to excellent growth for students at all achievement levels (Sanders and Rivers, 1996).

This is consistent with research from Koedel, Mihaly, & Rockoff (2015) in which they document the research showing substantial variation in teacher performance in terms of value-added scores. The argument of bias (in terms of overstating teacher impact) with value-added results was explored and found to be small when all factors were taken into account. Their research also explored various value-added models including one vs. two step models and an in-depth analysis of what we know about teacher value-added information. This included the importance of the research to how the data was positively but imperfectly correlated across
subjects and even the policy implications of using value-added data to make personnel decisions (Koedel, Mihaly, & Rockoff, 2015).

When examining the characteristics of highly effective teachers, research seems to indicate some surprising results. General academic ability or intelligence does not seem to have any bearing on teacher effectiveness though having a higher verbal ability does seem to lead to a higher teacher effectiveness. Knowledge of teaching and learning shows a positive relationship with teacher effectiveness while knowledge of subject matter does not seem to have any bearing on it. Teaching experience is an interesting characteristic as experience seems to matter for the first five years and then any relationship disappears. The number of teachers who have a degree in the field they are teaching is a better predictor of student achievement than experience or education level of teachers. However, teachers in states where there is a high rigor in terms of standards for the teaching profession have the highest performing students. (Darling-Hammond, 2000). This research was collected from data that was gathered by a 50-state survey and a whole host of case studies completed.

**Implications of Using Value-added for Teacher Evaluations**

As the educational climate has shifted away from proficiency towards ensuring all students succeed, value-added scores have become the new basis for determining teacher effectiveness. In Ohio, student scores, called student growth measures, can make up to 50% of a teacher's evaluation. For teachers who teach grades and subjects that provide value-added reports, this puts tremendous pressure on teachers to ensure their students perform well. As Braun's (2005) primer on value-added models (VAMs) notes, this is a positive step as it moves the discussion surrounding teacher quality to where it should be: focused on increasing student learning as the primary objective of teaching.
One of the possible implications of using VAM is that certain variables become far more important. For example, class make-up could become more complicated now due to value-added. In an ideal world, students would be randomly assigned to teachers so that no inherent bias could be found. This, however, cannot be guaranteed since districts could choose to honor parent requests, teacher seniority, special education assignments, etc. Since the ways in which students are matched to teachers may be related to students' potential, this can lead to a teacher being inappropriately credited or penalized for students' scores. Since this is a theoretical argument, it can be challenge to control for it in practice (Braun, 2005). This can also be seen in teacher concerns of having gifted students clustered in their classroom. Another variable to consider is measurement error on achievement tests due to factors such as familiarity with tests, the selection of specific questions, etc. that can skew students' scores (Hanushek & Rivkin, 2010). Student variability in terms of intelligence, home life, peer groups, etc. also have been shown to cause the validity of results to be questioned in terms of student achievement. A correlation of 0.4 to 0.7 have been shown for these types of uncontrollable factors in terms of achievement scores (Soar, Medley and Coker, 1983).

There are appropriate uses of VAM results that have an impact on teachers. Value-added results can be used as a screening strategy to determine what teachers are in the greatest need of professional development (Braun, 2005). Due to the limited amount of time and high number of teachers assigned to administrators, having a screener could be quite beneficial in providing the appropriate offerings to staff.

Teacher Evaluation and Models

When examining teacher evaluation, it is important to keep an unbiased approach to discussing teacher effectiveness. As such, no one measure should be weighted too heavily such that other methods are discredited. As Braun (2005) outlined, value-added results should not be
the sole or primary basis for making important decisions about teachers including retention or bonuses. Furthermore, they should not stop the examination of other appropriate measures of evaluation such as observation, portfolios, student surveys, etc.

There are different VAMs that are currently being used in various states and districts. As outlined in Braun's (2005) report, the most common models include EVAAS (the Educational Value-Added Assessment System), DVAAS (the Dallas Value-Added Accountability System) and REACH (Rate of Expected Academic Change). They each have slightly different models and methods for computing their scores. Ohio currently uses the EVAAS system that expresses a student score based on the sum of components such as district average for that grade and year, teacher effect and systematic and unsystematic variations. In this case, the teacher effect “is based on many different elements, including the growth in learning (as measured by an increase in test scores) of the students in the teacher’s classes over a number of years, adjusted for the effects of previous teachers of those students; the growth of the teacher’s students in subsequent years; and the achievements of those students in different subjects over a number of years, all appropriately adjusted for the contributions of those students’ other teachers (Braun, 2005, p. 12).”

**Teacher Perceptions of Gifted Students**

There is some research into teacher perceptions of giftedness, gifted education and working with gifted students. A study by McCoach & Siegle (2007) examined the attitudes of 262 teachers who participated in their study regarding their view on gifted education. Ironically, they found that teachers' attitudes about gifted did not change with training about gifted education. Special education teachers, however, were less supportive of acceleration and gifted education practices though (McCoach & Siegle, 2007). This raises some questions about professional development programs and their effectiveness in changing teacher perceptions of
working with gifted students. This conflicts with work completed by Copenhaver & McIntyre (1992) that found there were perceptual differences based on the grade level taught and what, if any, training teachers had undergone regarding gifted education (more training meant better perceptions). The overriding theme was that more targeted PD and work directly with students would be beneficial in improving teacher perceptions (Copenhaver & McIntyre, 1992).

Additional literature focused on possible differences in teacher perceptions based on their students' gender. A convenience sample of 210 schools in 30 states completed by Siegle & Reis (1998) focused on the gender differences in terms of ability, work ethic and quality of work. They found that teachers rated female students higher in terms of effort and quality of work while both genders rated similarly on academic areas except language arts where females rated higher. No information was provided that explored whether there were noticeable differences between male and female staff in terms of their perceptions (Siegle & Reis, 1998). In contrast, a study completed by Busse, Dahme, Wagner, & Wieczerkowski (1986) showed that male teachers were more likely to select male students and female teachers female students when asked if they had ever taught a highly gifted student. There was also a relationship by subject area with male students being disproportionally selected in areas like math and science while females were overly present in language arts (Busse, Dahme, Wagner, & Wieczerkowski, 1986).

**Gifted Service Models**

Since gifted education is a state issue, the Ohio Administrative Code organizes approved types of gifted service models for districts to use. These are broken down into two categories, those directly involving a gifted intervention specialist (GIS) and those that do not. Models involving a GIS include resource room pull-outs, replacement courses, co-teaching/inclusion and more. Those that do not directly involve a GIS include Advanced Placement (AP) classes, acceleration (subject or whole grade advancement), honors classes, and more. Based on the type
of model, there are specifics requirements that have to be put in place (Ohio Administrative Code 3301-51-15, 2008).

In terms of effective gifted service models, research done by Eric Calvert (2013) at Northwestern's Center for Talent Development highlights the effect size of some of the more common service models. The pull-out/resource room model is mildly effective (.3 to .6), but can be challenging to implement as they require a high level of staffing and can lump all types of gifted students together. Readiness or cluster grouping (up to .62.) can be very effective if ongoing assessment and regrouping occurs. Acceleration (whole or subject specific) tends to show the most growth (.5 to .8) when properly implement. In comparison, differentiation in the regular classroom has a small effect (less than .3) without having structural supports in place such as resources and training for staff (Calvert, 2013).

Summary, Discussion, Conclusion

The literature review revealed that many studies have been done to examine teacher characteristics that lead to student achievement. This should be no surprise as teacher effectiveness has long been an area of research. The challenge is the majority of the research is rather dated and needs to be updated going forward. Much of the literature available is from the 1980s or 1990s and the educational climate has changed dramatically since then. Initiatives such as NCLB, Race to the Top, and now the Every Student Succeeds Act (ESSA) have altered the focus in education slightly from proficiency to innovative solutions to close gaps to the need for teacher accountability with student growth. Value-added data can be used for some broad understandings in terms of teaching, teacher effectiveness, and how these results can be used for evaluating staffing.

For gifted students, the findings show a possible concern about the ceiling effect based on the type of test administered. If a norm-referenced test is given, the ceiling effect seems to be
non-impactful. In many cases, the use of above-level testing is commonly used to give a more accurate picture of a student's true abilities. This is especially true in a linear subject such as mathematics where acceleration is more common than other subjects.

There does seem to be a gap in the literature in terms of teacher perceptions of having gifted students in their class when examining their value-added scores. It remains to be seen if a teacher's perception impacts their students' performance on standardized tests. With the dramatic shift in Ohio to using student growth measures on the Ohio Teacher Evaluation System, this seems a prime opportunity to examine these types of questions.

**Implications and Suggestions for Future Research**

The literature suggests that much research has been done regarding value-added models and how to properly interpret results. The time seems to be right for more detailed research regarding value-added results and the impact on the new teacher evaluation framework in Ohio. More specifically, how gifted students' test scores are impacting their teacher's evaluation. With up to 50% of a teacher's evaluation based on student growth measures, the climate is right for an examination into the impact of scores on teacher's evaluations and the impact this is having on teacher perceptions of having gifted students in their class.
CHAPTER III. METHODOLOGY

The purpose of this research was to gather data and teacher perceptions on an ever growing area of importance in the state: gifted students, their value-added scores and how teacher evaluations are impacted. At a basic level, districts are seeking to meet the state indicators for their gifted students. The newest version of the state report card includes a gifted indicator which includes 3 components: performance index, value-added and input points. The performance index is based on the percentage of gifted students scoring at each categorical level (basic, limited, proficient, accelerated and advanced). The value-added portion looks at how gifted students score compared to their predicted growth. The scale is organized on an A to F scale where A is well above expected growth, a C is meeting growth, and an F is well below expected growth. Input points are based on the percentage of students identified as gifted and the percentage of gifted identified students who are served with special consideration to minority and economically disadvantaged students. In order to meet the indicator, a district must have a performance index score of 115, a value-added score of a C, and 40 gifted input points (Ohio Department of Education, 2016). This includes showing growth with their value-added scores.

At a deeper level, concerns over showing growth with these students could be impacting teacher perceptions of having gifted students in their class and this chapter will address the research method. First, the research design of the study was explored including the specifics of the research. Next, the participants of the study were addressed. After that, the instrumentation and data sources used to collect the respondents' information were thoroughly explained. Then, the data collection sources and the research questions have been presented. Once the data was collected, the data analysis occurred according to the steps laid out in this chapter. Finally, the assumptions, limitations and delimitations for the study were stated.

Research Questions
1. How is a district's service model related to teachers' progress with their gifted students in terms of value-added math scores?

2. Are districts concerned with the ceiling effect (students topping out by hitting a 99% on a test) for their gifted students, and if so, how are they attempting to minimize its impact?

3. How are teacher perceptions of having gifted students in their class related to their value-added math scores?

4. How are teacher perceptions of having gifted students in their class related to their OTES teacher evaluation ratings?

5. How do gifted teachers' value-added math scores compare to their OTES teacher rating?

**Research Design**

In designing a study to address these questions, a quantitative survey research study was the most logical choice. Data was collected with a survey from a convenience sample and was then analyzed. The survey collected data on teachers' value-added math scores, teacher and district perceptions on having gifted students in their class, service model used for gifted services, and the teachers' OTES teacher evaluation rating. The research methodology included both inferential and descriptive statistics. The relationship between the variables was key to speaking to which types of service models and teacher perceptions impacted student achievement if they did at all.

**Participants**

The participants sampled for this research were selected from the population of public school teachers in Ohio who teach gifted students. These thirty five participants were math teachers in grades 4-8 who have value-added data for identified gifted students in their classes.
In order to be a gifted education teacher, a teacher needs to have the gifted endorsement for their license. In today's educational world, that is accomplished by completing five graduate classes and an internship. The true number of gifted education teachers in Ohio is hard to quantify as there are other methods for serving gifted students without using qualified teachers. This sample of participants was both a convenience and purposive sample. It was a convenience sample as many of the respondents are people whom have information readily available with familiar routines and policies. It was also a purposive sample as individuals with a specific criteria were selected. For example, math teachers who had identified gifted students but no value-added data were excluded from this survey as they did not meet all the criteria.

The characteristics and size of the sample varied based on those who completed the survey. The participants sampled included all teachers who met the requirements previously stated from any public school setting in Ohio. The goal was to get a diverse sample of respondents including a variety of ethnic and economic situations. A large sample of 300 respondents was desired in order to get a variety of service and educational settings for comparative purposes.

A key contribution to this research study was the distribution of the study to the Ohio Association of Gifted Children's mailing list. OAGC's stated mission is “to promote and support the development of gifted students through dissemination of information, advocacy on their behalf, encouragement of affiliate organizations, and to promote research and education for gifted children (OAGC, 2017).” OAGC was founded in 1952 and its membership includes teachers, coordinators, administrators and parents. The organization was reorganized in 2004 to include separate divisions to promote the views of each of the different stakeholders (OAGC, 2017).
For this study, a large spike in respondents was seen when OAGC emailed the survey out to their membership on January 17, 2017. 17 respondents completed the survey from December 17, 2016 to January 16, 2017. On January 17, 7 survey responses were logged along with 10 more over the next eight days. This seems to indicate that OAGC respondents contributed to a large percentage of the respondents. This can be seen as a positive in the sense that these respondents are likely to have a working knowledge of gifted education in Ohio though they may also be a source of bias in the study.

All manner of ethical considerations to protect the participants were undertaken including maintaining the confidentiality of the respondents. The survey did not collect any identifiable information such as school name, teacher name or email address. The survey was emailed to building contacts listed in the state OEDS system along with shared with the Ohio Association of Gifted Children's mailing list. The unidentifiable information that was examined was kept on a secure computer for the minimally required three years. All aspects of the university's IRB process were followed to protect participants including the survey serving as implied consent and no benefits being accrued to those who respond. Any discomfort or inconvenience to the respondent derived only from the amount of time taken to complete the survey. The decision whether or not to participate did not prejudice any future relationships with The University of Findlay.

**Instrumentation and Data Sources**

In order to get the clearest picture possible of teacher perceptions, a survey was created by the researcher to collect teacher responses. This survey was constructed in consultation with the dissertation chair to ensure the questions were as straightforward and unbiased as possible. The survey was a Google Form that was very user friendly and provided an easy interface for respondents. The survey questions were as follows:
1. **Did you have a value-added score for the 2015-16 school year?** This was a yes or no choice and was essential as only teachers who had a value-added score could be included. For value-added, a teacher could score at one of five levels: dark green, light green, yellow, orange and red. Dark green means there is significant evidence that the school's students made more progress than the Growth Standard; light green means there is moderate evidence that the school's students made more progress than the Growth Standard; yellow is evidence that the school's students made progress similar to the Growth Standard, orange is moderate evidence that the school's students made less progress than the Growth Standard; red is significant evidence that the school's students made less progress than the Growth Standard. Specific values for each level vary by test.

2. **What was your value-added score for your math students for the 2015-16 school year?** A text box was provided for them to indicate their score for math students only as that is the specific population being examined.

3. **What type of service model(s) did you work in for the 2015-16 school year?** A drop down menu was provided to ensure common language was used for service models. Service models included options in the regular classroom without a gifted intervention specialist such as subject acceleration, cluster grouping (a group of gifted students placed together in one class) or honors classes along with models that include a gifted intervention specialist such as a pull-out class, single subject replacement course or inclusionary model. These were selected as they are the most often used service models as reported by the state.

4. **Do you feel you need to show growth with gifted students value-added scores?** This question was used to see the teacher's perception about showing growth.
5. **What do you think about your ability to help gifted students show growth with their value-added?** This question was an open text box and was included to gather any teacher insight into showing growth.

6. **Is your district concerned about the possibility of the ceiling effect (students topping out the state assessment) for your gifted students?** A drop down menu was provided to keep the choices the same.

7. **What method(s) is/are your district using to help your gifted students show progress?** A drop down menu was provided to ensure common language was used for methods that districts used.

8. **What was your student growth measures OTES rating for the 2015-16 school year?** A drop down menu was included with the five rating levels.

9. **What was your summative OTES rating for the 2015-16 school year?** A drop down menu was provided with the four rating levels.

**Data Collection Sources**

The survey was emailed from the researcher's University email account to educational contacts taken from the OEDS database in September. This list included superintendents, building principals and vocational centers for public schools. In addition, the survey was shared with the Ohio Association of Gifted Children's mailing list. The survey stayed open for two months through mid-November in order to provide enough time to gather responses before shutting it down before Thanksgiving. Teacher level value-added reports were released in September so this survey was sent out at a critical time when teachers' scores were fresh in their mind. Reminder emails were sent out one month in and then one week before the survey ended. These reminders provided the needed reminders to gather as many responses as possible.
The survey data was collected by the researcher throughout the fall and filled directly into a Google spreadsheet. This information was kept secure and remained open to respondents until the responses stop being collected in the middle of November. When the survey closed, the responses were sorted and put into a new spreadsheet. This was the spreadsheet used for data analysis.

**Data Analysis**

The first research question was focused on the impact a district's service model has on gifted students' value-added scores. The independent variable was the service model used by the district and the dependent variable was the teacher's value-added math scores. The survey question for the service model was a drop down menu while the teacher's value-added math scores was a text box. The drop down selections for service model included the following options: resource room, inclusion, replacement course (teacher of record), honors class or other. An ANOVA was run to examine the relationship between the service models and the resulting student score. This test was run using Excel.

The second research question was focused on how districts have minimized the ceiling effect with their gifted students if they feel it was an issue at all. The independent variable was the district perceptions of the ceiling effect and the dependent variable was the methods districts were using to minimize the impact of the ceiling effect. The survey question for the district perceptions was a multiple choice selection while the districts methods score was a check box. The multiple choice selections for district perceptions included the following options: yes, my district is very concerned about the ceiling effect; yes, my district is concerned about the ceiling effect; neither yes or no; no, my district is not concerned about the ceiling effect; no, my district is not worried at all about the ceiling effect. Only the methods provided by teachers who reported very concerned or concerned about the ceiling effect were examined for the comparison.
A descriptive comparison was run to compare the percentages for those respondents (concerned and very concerned) using different methods and the percentages of the way they perceived the impact of the ceiling effect. The various percentages were reported during the comparison. This test was run using Excel.

The next research question examined if teachers' perceptions of having gifted students in their class impacted their students' value-added math scores. The independent variable was the teachers' perceptions and the dependent variable was the teacher's value-added math scores. The survey question for the teacher perceptions was a multiple choice selection while the teacher's value-added math score was a text box. The multiple choice selections for teacher perceptions included the following options: I am confident that my gifted students can show growth, I feel it is likely that my gifted students can show growth, I am unsure if my gifted students can show growth, I feel it is unlikely that my gifted students can show growth, I am not confident my gifted students can show growth. An ANOVA will be run to examine the relationship between the teachers' perceptions and the resulting student score. The independent variable of teachers' perceptions was compared to student scores to examine if a teacher with a perception that they can show growth had student's who scored higher. This test was run using Excel.

The next research question examined if teachers' perceptions of having gifted students in their class impacted their OTES teacher evaluation rating. The independent variable was the teachers' perceptions and the dependent variable was their OTES teacher evaluation rating. The survey question for the teacher perceptions was a multiple choice selection while the teacher's OTES teacher evaluation rating was a text box. The multiple choice selections for teacher perceptions included the following options: I am confident that my gifted students can show growth, I feel it is likely that my gifted students can show growth, I am unsure if my gifted students can show growth, I feel it is unlikely that my gifted students can show growth, I am not confident my gifted students can show growth, I feel it is unlikely that my gifted students can show growth, I am not
confident my gifted students can show growth. An ANOVA will be run to examine the relationship between the teachers' perceptions and the teacher OTES rating. The independent variable of teachers' perceptions was compared to their OTES teacher evaluation rating to examine if a teacher with a perception that they can show growth had a higher OTES teacher evaluation rating. This test was run using Excel.

The final research question compared the teacher's value-added scores with their OTES teacher rating. The independent variable was the teacher's value-added score and the dependent variable was the OTES teacher rating. The survey question for the teacher's value-added math score was a text box as was the OTES teacher rating. The OTES teacher rating will be translated into a numerical value as follows: ineffective- 1, developing- 2, skilled- 3, and accomplished- 4. A t-test was run to compare the averages between the two scores. Close attention was paid to determine if those scoring higher on their OTES teacher rating score had a higher value-added score. The percentages of those scoring at each OTES level and their corresponding average value-added score were reported. This test was run using Excel.
Table 3. *Research Question Summary*

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Independent Variable</th>
<th>Dependent Variable</th>
<th>Statistical Test Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>How is a district’s service model related to teachers’ progress with their</td>
<td>District’s service model</td>
<td>Teachers’ value-added math scores</td>
<td>ANOVA</td>
</tr>
<tr>
<td>gifted students in terms of value-added math scores?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are districts concerned with the ceiling effect (students topping out by hitting</td>
<td>District perceptions on the ceiling effect</td>
<td>Methods districts are using to minimize</td>
<td>Descriptive Comparison</td>
</tr>
<tr>
<td>a 99% on a test) for their gifted students, and if so, how are they attempting</td>
<td></td>
<td>the impact of the ceiling effect</td>
<td></td>
</tr>
<tr>
<td>to minimize its impact?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How are teachers’ perceptions of having gifted students in their class related</td>
<td>Teachers’ perceptions</td>
<td>Teachers’ value-added math scores</td>
<td>ANOVA</td>
</tr>
<tr>
<td>to their value-added math scores?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How are teachers’ perceptions of having gifted students in their OTES teacher</td>
<td>Teachers’ perceptions</td>
<td>Teachers’ OTES teacher evaluation rating</td>
<td>ANOVA</td>
</tr>
<tr>
<td>evaluation ratings?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How do gifted teachers’ value-added math scores compare to their OTES teacher</td>
<td>Teachers’ value-added math scores</td>
<td>Teachers’ OTES teacher evaluation rating</td>
<td>T-Test</td>
</tr>
<tr>
<td>rating?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Assumptions**

The major assumption in this study was that the respondents completed the survey truthfully. Since it was an anonymous survey, there was no way to verify a respondent was being truthful so they had to be taken at their word. Since there were no consequences for the respondent based on their participation, they should have felt free to be open and honest in their responses. Another assumption was that the survey that was created was straightforward and not open to interpretation from the respondent. The questions were looking for specific answers rather than respondents open ended comments so the responses should have been accurate and applicable for the research.

Since the respondents were expected to have a working knowledge of gifted education, the assumption was made that they knew what terms such as high quality professional
development, service models and the ceiling effect meant. High quality professional
development has been ambiguous in recent years with the new operating standards seeking to
better define what it means. For this study, the assumption is that respondents understood high
quality professional development would be focused, ongoing support to improve instructional
practices for their gifted students. In terms of service models, professionals working in schools
would be providing gifted services in a variety of ways. The assumption was that they would be
aware of these models. Finally, the working assumption was also made that they would be
familiar with the ceiling effect and understand this referred to high ability students topping out
on assessments.

This research had some limitations based on the fact the state tests switched to a new
vendor each of the past two years and there were a variety of concerns with the results. Some
districts used paper/pencil versions of the test while others completed online assessments. There
were also a good deal of students in the state who opted out of testing the past two years which
removed some scores from teachers data. When collecting the teacher's value-added score in the
survey, these issues may pose a problem in the data. Since the research methodology was
focused on inferential statistics, there was no way to be completely sure that the values
calculated were correct since the population was not fully measured. In addition, the survey was
created by the researcher so there was the possibility the instrument might be not as reliable as
hoped.

There were a couple of delimitations with this research design. First, since gifted students
were a point of focus, only teachers who had identified gifted students in their class were
included in the research. Second, studying the ceiling effect was an essential part of this study so
only math teachers were included to collect information on districts who were math accelerating
students. Finally, value-added was a key component of the research so only teachers in grades 4-8 were respondents.
CHAPTER IV. RESULTS

This dissertation was a research study about teachers' perceptions of gifted students including if it was possible to show growth with their value-added scores and what, if any, impact their scores had on the teachers' summative OTES evaluation ratings. The study was focused on how OTES is impacting teachers' work with gifted math students. The student growth measures portion of teachers' evaluations are based on how students perform. Teachers' perceptions of the ceiling effect were collected along with what actions, if any, there district took to minimize the impact of the ceiling effect. Comparisons were run to look at their student growth measures in light of their perceptions about working with gifted students.

The study is based on the perceptions of Ohio public school math teachers who teach identified gifted students. A survey was created and distributed to teachers throughout the state who met the criteria mentioned above. The first, third, and fourth research questions were addressed through a one-way analysis of variance (ANOVA). Due to concerns over the small sample size, an Eta-Squared was run to determine the effect size. The second research question was completed using descriptive statistics to collect the number of respondents worried about the ceiling effect. The final research question was based on running a T-Test to compare teacher scores and their summative evaluation.

Characteristics of the Sample

Ohio public school gifted teachers who taught math were invited to participate in this study. The survey was distributed through membership organizations for teachers and coordinators of gifted students. In addition, the Ohio Association for Gifted Children (OAGC) included the survey in their monthly correspondence to their membership which includes teachers from across the state. Due to this distribution method, it is impossible to say how many
teachers received the survey. For the teachers who received this survey, 35 respondents completed and submitted a response online.

Among the 35 respondents who completed the survey created as a Google Form, 5 respondents were extracted as they did not have value-added data for the 2015-16 school year. The current study focused on the remaining 30 Ohio teachers who completed the survey. A further 5 respondents included a value-added score that was not in the requested numerical form. These included a question mark, the word average, a letter grade of F, not sure and n/a. These respondents were removed from the research questions that required a value-added score in numerical form though they were included in the other research questions.

The survey went live on December 19, 2016 and was included in emails to respondents that day. Some initial respondents returned their surveys on December 20 with a large break occurring from December 20 to January 1, 2017. The researcher assumed this was due to the holiday season. Responses continued throughout January with a large uptick on January 17 due to OAGC sending out the survey to their membership. The last entry was logged on February 7.

**Instrument Validity and Reliability**

The survey that was used was created in consultation with committee members and other education professionals. A Google Form was selected as the data collection method to disseminate the survey due to its ease of use, familiarity in the education community and lack of cost associated with its use. A few modifications were made based on feedback such as removing non-essential questions and clarifying what value-added score was being collected (index score). The survey was revised twice until it was pared back to collect just the essential information for this survey. Before going live, a fellow professional completed a sample survey to ensure the process worked properly.
Reliability was not a major concern as the survey was just collecting facts from respondents. There was only one survey item per construct so traditional reliability measures were not applicable in this study. In terms of validity, content validity was established by ensuring the survey collected the information it needed to in a consistent manner from respondent to respondent. The majority of the survey questions were multiple choice selections to ensure common responses by the respondents.

**Research Question 1**

The first research question examined the relationship between a teacher's value-added score and the service model they worked in. Specifically, does the service model the teacher works in impact their value-added score. The service models were presented in a list that included resource room, inclusion/co-teaching, replacement course (teacher of record), honors class, subject acceleration, regular classroom with high quality professional development or other. These are all service models taken from the state's operating standards for gifted education (Ohio Administrative Code 3301-51-15, 2008). The value-added score was their teacher index score they provided in the blank provided.

**Research Question 1 Answered**

The responses (n = 25) to question 1 were sorted based on their service model selected. The inclusion/co-teaching service model did not have any respondents so it was dropped from the calculations. A further two categories (resource room and “other” which gifted cluster classroom was entered) only had one response each so they were excluded as well. Of the remaining five service models, the replacement course (teacher of record) was the most popular response (n = 9). From there, it went to regular classroom with high quality professional development (n = 6), subject acceleration (n = 5) and honors class (n = 3). Table 4 details not only the number of respondents but also the mean and variance for each service model type
included. The highest mean (4.205) and lowest variance (2.106) both belong to the regular classroom with high quality professional development.

Table 4. *Number of Respondents, Mean Value-Added Score and Variance by Service Model Type*

<table>
<thead>
<tr>
<th>Service Model</th>
<th>Number of Respondents</th>
<th>Mean Value-Added Score</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replacement Course (Teacher of Record)</td>
<td>9</td>
<td>1.859</td>
<td>3.983</td>
</tr>
<tr>
<td>Regular Classroom with HQPD</td>
<td>6</td>
<td>4.205</td>
<td>2.106</td>
</tr>
<tr>
<td>Subject Acceleration</td>
<td>5</td>
<td>2.586</td>
<td>5.523</td>
</tr>
<tr>
<td>Honors Class</td>
<td>3</td>
<td>2.637</td>
<td>9.834</td>
</tr>
</tbody>
</table>

An Analysis of Variance (ANOVA) was run to see if there was any significant difference among value-added scores by service model. The alpha value was set at 0.05 and the P-value was 0.244 which would tend to mean there is a failure to reject the null hypothesis. This is confirmed by having an F value (1.511) that is lower than the F critical value (3.127). However, when investigating further the concern is the small sample size is influencing the effect size. At this point, an Eta-Squared calculation was run that compared the Sum of Squares (SS) between the groups (20.013) and the total SS (103.914). This provided an Eta-Squared value of 0.193 which is a significant large effect size.

Based on the statistics, it can be said there was a large effect size which means that if the same effect was seen in a large sample, there would be a statistically significant difference in teachers' value-added scores based on the service model they teach in. In this study, teachers in a regular classroom with high quality professional development had this highest average value-added score. This is rather surprising when compared to other work such as that of Eric Calvert where subject acceleration had a higher effect size than differentiation in the regular classroom on student achievement (Calvert, 2013).
**Research Question 2**

The second research question explored the connection between a district's perception of the ceiling effect (seen as students topping out on state assessments) and what, if any, methods they were using to try and minimize the impact of this phenomenon. The question for a district's perception provided five choices for the respondent to select from: yes, my district is very concerned about the ceiling effect; yes, my district is concerned about the ceiling effect; neither yes or no; no, my district is not concerned about the ceiling effect and no, my district is not worried at all about the ceiling effect. For the methods districts are using, there were three choices provided (tracking gifted students' value-added scores, math accelerating high performing students to the next grade level and pursuing professional development aimed at high performing students) along with an other box.

**Research Question 2 Answered**

The responses (n = 30) to the second question were separated into two groups: those who were concerned or very concerned about the ceiling effect and those who were neutral or not concerned or worried about the ceiling effect. For the purposes of this study, only those who answered as concerned or very concerned (n = 13) were evaluated further to see what methods they were using to mitigate the impact of the ceiling effect as seen in Table 5. Only two respondents indicated their district was very concerned about the ceiling effect. Both of them identified that they were using the same methods to adjust for the impact of the ceiling effect: tracking gifted students' value-added scores, math accelerating high performing students to the next grade level, and pursuing professional development aimed at high performing students.

The number of districts who indicated they were concerned about the impact of the ceiling effect (n = 11) was more varied than the two who identified as very concerned. The number one responses by far (n = 8) was that they were math accelerating high performing
students to the next grade level. The other two provided choices, tracking gifted students' value-added scores \((n = 2)\) and pursuing professional development aimed at high performing students \((n = 2)\), were both selected by two respondents as being used in their districts. Two other methods were identified by respondents in the “other” field: extensions tied to above grade level standards \((n = 1)\) and looking for a rigorous curriculum where student thinking, reasoning and problem solving build an understanding \((n = 1)\).

Table 5. *Number of Respondents Concerned or Very Concerned About the Ceiling Effect and the Methods Used to Mitigate the Impact*

<table>
<thead>
<tr>
<th>Yes, my district is very concerned about the ceiling effect ((N = 2))</th>
<th>Yes, my district is concerned about the ceiling effect ((N = 11))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Method</strong></td>
<td><strong>Method</strong></td>
</tr>
<tr>
<td>Tracking gifted students' value-added scores</td>
<td>Tracking gifted students' value-added scores</td>
</tr>
<tr>
<td>(N = 2, 100%)</td>
<td>(N = 2, 18.2%)</td>
</tr>
<tr>
<td>Math accelerating high performing students to the next grade level</td>
<td>Math accelerating high performing students to the next grade level</td>
</tr>
<tr>
<td>(N = 2, 100%)</td>
<td>(N = 8, 72.7%)</td>
</tr>
<tr>
<td>Pursuing professional development aimed at high performing students</td>
<td>Pursuing professional development aimed at high performing students</td>
</tr>
<tr>
<td>(N = 2, 100%)</td>
<td>(N = 2, 18.2%)</td>
</tr>
<tr>
<td>Other</td>
<td>Other: extensions tied to above grade level standards</td>
</tr>
<tr>
<td>(N = 0, 0%)</td>
<td>(N = 1, 9.1%)</td>
</tr>
<tr>
<td>Other</td>
<td>Other: looking for a rigorous curriculum where student thinking, reasoning and problem solving build an understanding</td>
</tr>
<tr>
<td>(N = 0, 0%)</td>
<td>(N = 1, 9.1%)</td>
</tr>
</tbody>
</table>

Based on the data, districts who are either concerned or very concerned about the ceiling effect are using the same methods listed to alleviate their concerns. All the districts who identified as very concerned about the ceiling effect selected all three provided options as the methods they were using. The most common method for either type of district is to subject
accelerate high performing math students. The remaining methods were very similar in terms of the likelihood that districts are using them.

**Research Question 3**

The third research question compared teachers' perceptions of showing value-added growth with their students and the teachers' value-added scores. For the teachers' perceptions of showing value-added growth with their gifted students, five choices were provided. These included “I am confident that my gifted students can show growth”, “I feel it is likely that my gifted students can show growth”, “I am unsure if my gifted students can show growth”, “I feel it is unlikely that my gifted students can show growth” and “I am not confident my gifted students can show growth.” The value-added score was their teacher index score they provided in the blank provided.

**Research Question 3 Answered**

The responses (n = 25) to question 3 were filtered based on the teachers' perceptions of their ability to help their gifted students show value-added growth. The options “I feel it is unlikely that my gifted students can show growth” and “I am not confident my gifted students can show growth” had zero respondents and were dropped from the calculations. Of the remaining three choices, all three had a similar number of responses. Table 6 details not only the number of respondents but also the mean value-added score and variance for each perception phrase included. The highest number of responses was “I feel it is likely that my gifted students can show growth” (n = 10) followed by “I am confident that my gifted students can show growth” (n = 9) and then “I am unsure if my gifted students can show growth” (n = 6).
Table 6. Number of Respondents, Mean Value-Added Score and Variance Based on Teacher’s Perception

<table>
<thead>
<tr>
<th>Teachers’ Perceptions</th>
<th>Number of Respondents</th>
<th>Mean Value-Added Score</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel it is likely that my gifted students can show growth</td>
<td>10</td>
<td>3.139</td>
<td>6.595</td>
</tr>
<tr>
<td>I am confident that my gifted students can show growth</td>
<td>9</td>
<td>3.133</td>
<td>3.768</td>
</tr>
<tr>
<td>I am unsure if my gifted students can show growth</td>
<td>6</td>
<td>1.897</td>
<td>3.433</td>
</tr>
</tbody>
</table>

An Analysis of Variance (ANOVA) was run to see if there was any significant difference among value-added scores by service model. The alpha value was set at 0.05 and the P-value was 0.497 which would tend to mean there is a failure to reject the null hypothesis. This is confirmed by having an F value (0.723) that is lower than the F critical value (3.443). In order to ensure the small sample size was not influencing the effect size, further calculations were completed. An Eta-Squared calculation was run that compared the Sum of Squares (SS) between the groups (7.008) and the total SS (113.663). This provided an Eta-Squared value of 0.062 which reveals no significant effect size.

Based on these calculations, it can be said there is no significant difference in teachers’ value-added scores based on their perception of showing growth with their gifted students.

**Research Question 4**

Research question 4 explored the connection between teachers’ perceptions of showing value-added growth with their students and the teachers’ final summative OTES evaluation rating. For the teachers' perceptions of showing value-added growth with their gifted students, five choices were provided. These included “I am confident that my gifted students can show..."
growth”, “I feel it is likely that my gifted students can show growth”, “I am unsure if my gifted students can show growth”, “I feel it is unlikely that my gifted students can show growth” and “I am not confident my gifted students can show growth.” The teachers' summative OTES evaluation rating was a choice of the four rating levels: accomplished, skilled, developing and ineffective.

**Research Question 4 Answered**

The responses (n = 30) to question 4 were sorted based on the teachers' perceptions of their ability to help their gifted students show value-added growth. The phrase “I feel it is unlikely that my gifted students can show growth” had zero respondents and was dropped from the data. Another choice, “I am not confident my gifted students can show growth”, only had one response and was not included in the calculations. Of the remaining three choices, all three had a similar number of responses. Table 7 details not only the number of respondents but also the mean summative OTES rating and variance for each perception phrase included. The highest number of responses was “I feel it is likely that my gifted students can show growth” (n = 12) followed by “I am confident that my gifted students can show growth” (n = 10) and then “I am unsure if my gifted students can show growth” (n = 7).

For the teachers' summative OTES value, they were asked to select their final rating from a list of the four provided: accomplished, skilled, developing and ineffective. Each rating was assigned a corresponding value with accomplished equaling 4, skilled = 3, developing = 2 and ineffective = 1.
Table 7. \textit{Number of Respondents, Mean Summative OTES Evaluation and Variance Based on Teacher's Perception}

<table>
<thead>
<tr>
<th>Teachers' Perceptions</th>
<th>Number of Respondents</th>
<th>Mean Summative OTES Evaluation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel it is likely that my gifted students can show growth</td>
<td>12</td>
<td>3.5</td>
<td>0.273</td>
</tr>
<tr>
<td>I am confident that my gifted students can show growth</td>
<td>10</td>
<td>3.6</td>
<td>0.267</td>
</tr>
<tr>
<td>I am unsure if my gifted students can show growth</td>
<td>7</td>
<td>3.429</td>
<td>0.286</td>
</tr>
</tbody>
</table>

An Analysis of Variance (ANOVA) was run to see if there was any significant difference among teachers' summative OTES evaluation scores by teachers' perceptions. The alpha value was set at 0.05 and the P-value was 1.0 which would tend to mean there is a failure to reject the null hypothesis. This is confirmed by having an F value (0.232) that is lower than the F critical value (1.846). In order to ensure the small sample size was not influencing the effect size, further calculations were completed. An Eta-Squared calculation was run that compared the Sum of Squares (SS) between the groups (0.127) and the total SS (7.241). This provided an Eta-Squared value of 0.018 which reveals no significant effect size.

Based on these calculations, it can be said there is no significant difference in teachers' summative OTES evaluation rating based on their perception of showing growth with their gifted students.

\textbf{Research Question 5}

The last research question compared the teachers' final summative OTES evaluation rating with their value-added score. For the teachers' summative OTES evaluation rating, the
four levels of accomplished, skilled, developing, and ineffective were possibilities. The value-added score was their teacher index score they provided in the blank provided.

**Research Question 5 Answered**

The responses (n = 25) to question 5 were split into two groups based on the teachers' final summative OTES evaluation rating. All the respondents indicated they were rated either accomplished or skilled on their final rating. Table 8 details not only the number of respondents but also the mean value-added score and variance for each summative OTES evaluation rating included. The highest number of responses were those who earned the accomplished rating (n = 13) followed by those who were rated as skilled (n = 12).

<table>
<thead>
<tr>
<th>Teachers' Summative OTES Evaluation Rating</th>
<th>Number of Respondents</th>
<th>Mean Value-Added Score</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accomplished</td>
<td>13</td>
<td>3.055</td>
<td>5.438</td>
</tr>
<tr>
<td>Skilled</td>
<td>12</td>
<td>2.605</td>
<td>4.286</td>
</tr>
<tr>
<td>Developing</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Ineffective</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

A T-test was run to see if there was any significant difference among teachers' value-added scores based on their summative OTES evaluation rating. The alpha value was set at 0.05 and the P-value was 0.372 which means there is a failure to reject the null hypothesis. Cohen's D provided a value of 0.193 which is a small effect size.

Based on these calculations, it can be said there is no significant difference in teachers' value-added scores based on their summative OTES evaluation rating. This appears to be an
issue for future research, as student growth measure scores such as value-added is factored into the final summative rating for OTES.

**Summary**

This research study sought to look into whether teachers' perceptions of working with and showing value-added growth with gifted students had an impact on teachers' value-added scores and teacher evaluation ratings. The first research question examined how a district's service model related to teachers' progress with their gifted students in terms of value-added math scores. There was a significant effect size found indicating the service model a teacher taught in impacted their value-added score. The highest mean value-added score was found with teachers who worked in the regular classroom with high quality professional development setting.

The second research question delved into whether districts were concerned with the ceiling effect (students topping out by hitting a 99% on a test) for their gifted students, and if so, how they attempted to minimize its impact. Almost half of the respondents (43.3%) indicated their district was either very concerned or concerned about the impact of the ceiling effect with their gifted students. The most common method used to mitigate this concern was by math accelerating students to the next grade level.

The third research question looked at how teacher perceptions of having gifted students in their class related to their value-added math scores. When looking at teachers' perceptions of showing growth with gifted students, no significant difference was found in teachers' value-added scores regardless of their perceptions. The fourth research question investigated how teacher perceptions of having gifted students in their class related to their OTES teacher evaluation ratings. No significant difference was found in teachers' summative OTES evaluation ratings based on their perceptions of showing growth with their gifted students. Finally, the fifth and final research question asked how gifted teachers' value-added math scores compared to their
OTES teacher rating. No noticeable difference was found between accomplished and skilled teachers in terms of their value-added scores.
CHAPTER V. CONCLUSIONS AND RECOMMENDATIONS

The purpose of this chapter is to summarize the findings of the research study and suggest recommendations for future policy, practice, and research opportunities. The first portion of this chapter will discuss an overview of the entire study with emphasis on the goals of the research and the methodology used to accomplish the quantitative analysis. A summary of the results for the five research questions will be described in detail. The second portion of the chapter will discuss policy and practice recommendations from the research findings. Lastly, recommendations for future research will be proposed as it relates to teachers' work with gifted students.

Review of the Study

The move to value-added models has been a positive trend in terms of providing opportunities to examine achievement for all students including high-ability ones. As noted by Braun (2005), focusing on value-added models shifts the focus of teaching onto student achievement. The purpose of this study was to examine whether a teacher's perceptions of showing growth with gifted students' value-added scores had an impact on the students' actual scores. In addition, the study focused on the service models used by districts and if they had an impact on student scores. This also led to an examination of whether a lack of student growth for high-ability students was negatively impacting their teachers' value-added scores and OTES rating.

This study took into account the thousands of teachers in Ohio that are now tied to student value-added scores. For the 2013-14 school year, this included over 4,600 5th grade math and reading teachers alone (SAS Institute, 2015). Data driven decision making and being proficiently fluent in student growth statistics are now a large part of a teacher's job. Furthermore, administrators have to make difficult personnel decisions including what students
each teachers is assigned which can be based in part on student value-added scores. Value-added is now a critical component of the public education system in Ohio.

A population of Ohio’s public school teachers who taught gifted math students full-time in grades 4-8 during the 2015-2016 school year were asked to participate in the online survey. The survey was distributed through membership organizations for teachers and coordinators of gifted students. In addition, the Ohio Association for Gifted Children (OAGC) included the survey in their monthly correspondence to their membership which includes teachers from across the state. From the total sampling population, a total of 35 usable responses to the instrument emerged. The instrument was administered via Google Forms, an online survey that was distributed through an email solicitation that provided a link to the questionnaire.

Prior to analyzing the research questions, the validity and reliability of the survey was determined. There was only one survey item per construct so traditional reliability measures were not applicable in this study. The majority of the survey questions were multiple choice selections to ensure common responses by the respondents. The survey questions were reviewed by experts in the field.

The first, third and fourth research questions were addressed through a one-way analysis of variance (ANOVA). Due to concerns over the small sample size, an Eta-Squared was run to determine the effect size. The second research question was completed using descriptive statistics to collect the number of respondents worried about the ceiling effect. Finally, the final research question was based on running a T-Test to compare teacher scores and their summative evaluation. The research questions were:

1. How is a district's service model related to teachers' progress with their gifted students in terms of value-added math scores?
2. Are districts concerned with the ceiling effect (students topping out by hitting a 99% on a test) for their gifted students, and if so, how are they attempting to minimize its impact?

3. How are teacher perceptions of having gifted students in their class related to their value-added math scores?

4. How are teacher perceptions of having gifted students in their class related to their OTES teacher evaluation ratings?

5. How do gifted teachers' value-added math scores compare to their OTES teacher rating?

Discussion

A recent focus in Ohio has been the implementation of the Ohio Teacher Evaluation System for public school teachers. There has been less of a focus on what, if any, impact gifted students are having on their teachers' OTES ratings and in turn, the impact teachers' perceptions are having on student value-added scores. The research questions were created to quantify the impact teachers' perceptions have on their gifted students in terms of their value-added growth and the teachers' OTES evaluations, whether the type of service model impacts scores, examining whether districts are concerned about the ceiling effect and how students' value-added scores are impacting teachers' evaluations.

Research Question 1.

The findings of this study indicate that there was a large effect size in gifted students' value-added scores based on the service model their district uses. Assuming this same effect was seen in a larger sample, there would be a statistically significant difference in teachers' value-added scores based on the service model they teach in. The Ohio Revised Code breaks service models down into two categories, those directly involving a gifted intervention specialist (GIS)
and those that do not. Models involving a GIS include resource room pull-outs, replacement courses, co-teaching/inclusion and more. Those that do not directly involve a GIS include Advanced Placement (AP) classes, acceleration (subject or whole grade advancement), honors classes, and more (Ohio Administrative Code 3301-51-15, 2008).

In this study, teachers in a regular classroom with high quality professional development had the highest average value-added score. The model with the most contact time with a gifted intervention specialist, replacement course/teacher of record, actually had the lowest average value-added score. These findings are rather surprising when compared to other work that has been done on the effectiveness of service models for gifted students. Studies such as that of Eric Calvert showed that subject acceleration had a higher effect size than differentiation in the regular classroom on student achievement (Calvert, 2013). It is important to note that the shortcomings listed in previous research (structural supports and teacher training) are accounted for in the service model selected in this study.

Due to the small sample size, the findings of this research need to be taken in the proper context. A replication of this study with a larger sample size would be needed to make more concrete assertions.

Research Question 2.

The second research question examined whether districts were concerned with the ceiling effect for their gifted students, and if so, how they were attempting to minimize it's impact. As outlined by Koedel & Betts (2010), the ceiling effect can be seen as the tendency for growth in a student’s test score to be less if their initial score is at the top end of the scores as the student has little room for growth given the difficulty level of the test. The findings of the study showed that districts who identified as either concerned or very concerned about the ceiling effect all used similar methods to minimize the impact on their students' growth. The districts who identified
that they were very concerned used the same three methods to adjust for the impact of the ceiling
effect: tracking gifted students' value-added scores, math accelerating high performing students
to the next grade level and pursuing professional development aimed at high performing
students. Research on countering the ceiling effect such as that done by Hansen (1992) has
pointed to above-level testing which aligns with acceleration.

Research such as that by Silverman (2012) documents that grade-level group
achievement tests have low ceilings. That would be the case in this study where teachers are
being asked about math achievement tests. However, the current literature is lacking for the
ceiling effect in terms of value-added scores and the impact on teacher performance. There is
research indicating that value-added estimates are generally immune to ceiling effects (Koedel &
Betts, 2010) especially when using norm-referenced tests rather than a proficiency test.

**Research Question 3.**

One of the main components of this study was to examine how teacher perceptions of
gifted students impact their value-added math scores. Teacher perceptions are well researched in
some areas such as gender bias (Siegle & Reis, 1998; Busse, Dahme, Wagner, & Wieczerkowski,
1986) but not as much in the area of gifted education. The two studies referenced (McCoach &
Siegle, 2007; Copenhaver & McIntyre, 1992) were at odds in their conclusions over teacher
perceptions of working with gifted students.

While the mean value-added scores for both “I feel it is likely that my gifted students can
show growth” and “I am confident that my gifted students can show growth” were higher than
the scores for those who answered “I am unsure if my gifted students can show growth” (3.139
and 3.133 compared to 1.897), no significant difference in teachers' value-added scores were
found based on their perception of showing growth with their gifted students. There was a large
variance found in all three groups that ranged from 6.595 for those who responded “I feel it is
likely that my gifted students can show growth” to 3.433 for those who answered “I am unsure if my gifted students can show growth”. An Eta-Squared calculation was run that compared the Sum of Squares (SS) between the groups (7.008) and the total SS (113.663) which provided an Eta-Squared squared value of 0.062 which revealed no significant effect size.

The results of this study aligned with the research that indicated that teacher perceptions of working with gifted students do not impact student achievement. These findings are important in determining the needs of future professional development. Targeted professional development can continue to be focused on the OTES process rather than needing to change teacher perceptions.

Research Question 4.

The fourth research question was related to the third in that it also focused on teacher perceptions of working with gifted students. For this question, the focus was on how teacher perceptions of having gifted students in their class related to their actual OTES teacher evaluation ratings. At the heart of the matter was the idea of would teachers with a positive perception of gifted student growth score higher on their OTES ratings than those with an unsure view. The smaller sample size was of concern here but the Eta-Squared calculation revealed no significant effect size.

Respondents from each of the three teacher perception categories listed all had similar mean summative OTES ratings. When everything is calculated, it can be said there is no significant difference in teachers' summative OTES evaluation rating based on their perception of showing growth with their gifted students. As with the results from research question 3, these results were a positive finding for the teacher evaluation process as it revealed that teacher perceptions do not impact teachers abilities to meet the rigorous demands of OTES.

Research Question 5.
The fifth and final research question examined how gifted teachers' value-added math scores compared to their OTES summative teacher rating. The original notion was that high scoring teachers would get a higher OTES summative rating. The OTES model includes a large component of value-added scores when determining the final rating. Each OTES rating was given a corresponding point total for the purposes of calculation: Ineffective-1, Developing-2, Skilled-3 and Accomplished-4. Part of the challenge for this question was that respondents only indicated they had received a rating of accomplished or skilled, no developing or ineffective responses were collected.

The t-test showed that the mean value-added scores were very similar for either rating. Based on these calculations, it can be said there is no significant difference in teachers' value-added scores based on their summative OTES evaluation rating. This appears to be an issue for future research though as student growth measures such as value-added are factored into the final summative rating for OTES.

The findings for this research question demonstrated that value-added scores were not dependent on summative OTES ratings. While student scores are included in the student growth measures component, the OTES rubric seemed to help teachers negate any potential difference by allowing them to reach accomplished or skilled. Since teachers have a much greater control of the teacher component of the OTES process, these findings spoke to the fact that teachers have learned the OTES process and how to effectively navigate it.

**Conclusion**

Ohio has seen a fair amount of change with the teacher evaluation process which has now made classroom teachers' value-added scores even more visible to all interested parties. At the same time, gifted students have come to the forefront with the creation of the gifted indicator which now means the growth shown by these students is important. Previously, only a small
percentage of teachers felt high ability students were a point of focus in terms of their test scores (Loveless, et al., 2008). To this end, the literature shows that teachers are now aware of the types of tests and how students are scoring (Koedel & Betts, 2010; U.S. Department of Education, 2008). In contrast, there is beginning to be a push towards above-level testing in response to concerns over the ceiling effect (Stanley, 1990; Van Tassel-Baska, 1986; Northwestern University: Center for Talent Development, 2015; Silverman, 2012; Hansen 1992). A new twist is the role computer adaptive testing is now playing in education (Davey, 2011; Gershon, 2005; Colwell, 2013).

The role and use of value-added scores in teacher evaluation is well documented in a variety of research. Research has been done to outline the variables in student scores and what items are within a teacher's control (Braun, 2005; Hanushek & Rivkin, 2010; Soar, Medley and Coker, 1983). Braun's (2005) work also outlines teacher evaluation models and how to best organize a teacher evaluation system.

The research into teacher perceptions of working with gifted students conflicts with some research (McCoach & Siegle, 2007) finding professional development had no impact on them while other research (Copenhaver & McIntyre, 1992) disagrees with this view. More consistent views on teacher perceptions are found in research on gender bias (Siegle & Reis, 1998; Busse, Dahme, Wagner, & Wieczerkowski, 1986).

This study significantly contributed to developing an understanding of teachers' perceptions of gifted students and the impact on their value-added scores. This study provided insight into how concerned districts are about the ceiling effect and what measures they are taking to reduce its' impact. The study also delved into the teacher evaluation system and the impact of students' value-added scores.
The service models districts used were found to be a predictor of students' value-added scores. All the districts who indicated they were concerned or very concerned about the ceiling effect were found to use common methods to alleviate it such as subject accelerating high performing math students. Teachers' perceptions of gifted students were found to have no significant impact on their value-added scores. Due to the variance in responses and small sample size, repeating the study with a larger sample could yield additional information that could better inform targeted teacher professional development.

There was also no significant difference in teachers' summative OTES evaluation rating based on their perception of showing growth with their gifted students. Finally, it can be said there is no significant difference in teachers' value-added scores based on their summative OTES evaluation rating. These findings revealed that the OTES process has become familiar to teachers as their work with the teacher rubric component had more impact on their evaluation than the value-added scores of their students. This is a perk as a teacher has far more control over their planning, instruction and assessment than they do over student value-added scores.

**Recommendations**

Since the service models districts used were found to be a predictor of students' value-added scores, districts should think through the way they will provide services to gifted students. This study revealed that exposing students to advanced content, whether through acceleration or honors classes, should have a positive impact on student growth. The best performing service model in terms of value-added scores included high quality professional development being implemented for the classroom teacher.

This study reveals that any district who expresses concern over the ceiling effect with their high-ability students can gain insight from other districts. Those districts who have expressed concern are using the same process to show growth with these students. First, they are
tracking student's value-added scores so the teachers know how their students have scored the previous year. Second, districts are using subject (math) acceleration to move high performing students on to the next grade level. This allows them to be exposed to more advanced content and show growth as the ceiling has been raised. Finally, districts are implementing target professional development for teachers to meet the growth needs for high-ability students.

This research revealed that teachers' perceptions of gifted students were found to have no significant impact on either students' value-added scores or the teachers' summative OTES evaluation rating. This is not all that unexpected since the bulk of respondents selected they felt it was likely their gifted students could show growth. Still, this study indicates there is no identified need to change teacher perceptions.

**Future Research Opportunities**

The implications of this study include the recognition that the service model a district uses impacts the growth of it's students. Further research is needed to gain a larger sample size to verify that the effect size seen in the research is consistent in a larger sample population. This research should inform district decision making when implementing gifted services. Since resources are scarce in education, this study can provide guidance into the best use for maximum student growth.

This study potentially holds important implications for districts which are worried about the value-added growth of their high ability students. There were common methods selected by all the schools who expressed concern over the ceiling effect which speaks to a possible field for further research. Additionally, the study could be replicated in a few years when more consistent findings are available from the new state testing that is being completed online.

The final recommendation for future research comes from the findings on student growth measures not significantly impacting the final summative rating for OTES. Since student growth
measures are a portion of the overall total, reason seems to indicate there should be a correlation of some type. Replicating this study with a larger sample size could yield important clarifications on this research question.
REFERENCES


Calvert, E. (2013, October). From high ability to high achievement: gifted service models that promote growth. Presentation given at the Ohio Association for Gifted Children's conference, Columbus, OH.


Institutional Review Board

Date:  November 11, 2016

To:      Jon Brasfield
Cc:      Brian T. Billings
RE:      Gifted Students Value-Added Scores and Teacher Evaluation

Project Expiration date:  November 11, 2017

The University of Findlay Institutional Review Board (IRB) has completed its review of your project utilizing human subjects and has granted authorization. This study has been approved for a period of one year only. The project has been assigned the number 1049.

In order to comply with UF policy and federal regulations, human subject research must be reviewed by the IRB on at least a yearly basis. If you have not completed your research within the year, it is the investigator’s responsibility to ensure that the Progress Report is completed and sent to the IRB in a timely fashion. The IRB needs to process the re-approval before the expiration date, which is printed above.

Understand that any proposed changes may not be implemented before IRB approval, in which case you must complete an Amendment/Modification Report.

Following the completion of the use of human subjects, the primary investigator must complete a Certificate of Compliance form indicating when and how many subjects were recruited for the study.

Please refer to the IRB guidelines for additional information. Please note that if any changes are made to the present study, you must notify the IRB immediately. Please include that number on any other documentation or correspondence regarding the study.

Thank you very much for your cooperation. If you have any questions, please feel free to contact IRB at (419) 434-4640 or email irb@findlay.edu.

Sincerely,

Susan W. Stevens, EdD., AT
Chair, Institutional Review Board
APPENDIX B INVITATION TO PARTICIPATE IN THE STUDY

Dear teachers,

You are invited to participate in a study of teacher perceptions of teaching gifted math students, value-added scores and Ohio Teacher Evaluation System (OTES) teacher evaluation ratings. We hope to learn without prejudice how you feel about teaching these students, what your value-added score has been and what, if any, impact these students have had on your OTES teacher rating. You were selected as a possible participant in this study because you were identified as an Ohio teacher who might teach a population of gifted math students. If you decide to participate, please complete the enclosed anonymous survey. Your return of this survey is implied consent. The survey is designed to elicit your perceptions about teaching gifted students. It will take about 15 minutes to complete the survey. No benefits accrue to you for answering the survey. Any discomfort or inconvenience to you derives only from the amount of time taken to complete the survey.

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will not be disclosed. The anonymous, unidentifiable data collected will be stored on a locked laptop in a locked office for the minimally required three years. Your decision whether or not to participate will not prejudice any future relationships with The University of Findlay. If you decide to participate, you are free to discontinue participation at any time without prejudice.

If you have any questions, please contact Brian T. Billings via email at billingsb@findlay.edu or the chairperson of The University of Findlay Institutional Review Board, IRB Chair via email at irb@findlay.edu.

Thank you for your time.

Sincerely,

Brian T. Billings
FIGURES

Figure 1. OTES Original Framework for Teacher Evaluation

Figure 2. OTES Alternate Framework for Teacher Evaluation
The evaluation factors are weighted as follows:

1. If a district chooses the original framework, the teacher performance measure and student growth measure shall be 50% each.
2. If a district chooses the alternative framework:
   - The teacher performance measure shall account for 50%;
   - The student academic growth measure shall account for 35%; and
   - The chosen alternative component(s) shall account for 15%.

Figure 3. OTES Final Summative Evaluation Rating for Teacher Evaluation Tables
### Table 1. Graduation Points Earned by Performance Level on End of Course Tests

<table>
<thead>
<tr>
<th>Performance Level</th>
<th>Graduation Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced</td>
<td>5</td>
</tr>
<tr>
<td>Accelerated</td>
<td>4</td>
</tr>
<tr>
<td>Proficient</td>
<td>3</td>
</tr>
<tr>
<td>Basic</td>
<td>2</td>
</tr>
<tr>
<td>Limited</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: Ohio Department of Education

### Table 2. Math Achievement Scores for Ohio Students for 2006-07

<table>
<thead>
<tr>
<th>Category</th>
<th>Ohio 4th Graders</th>
<th>Ohio 8th Graders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>State Data-</td>
<td>NAEP Data-</td>
</tr>
<tr>
<td></td>
<td>% Proficient</td>
<td>% Basic</td>
</tr>
<tr>
<td>All</td>
<td>76%</td>
<td>87%</td>
</tr>
<tr>
<td>White</td>
<td>82%</td>
<td>93%</td>
</tr>
<tr>
<td>Black</td>
<td>50%</td>
<td>67%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>61%</td>
<td>76%</td>
</tr>
<tr>
<td>Low Income</td>
<td>62%</td>
<td>75%</td>
</tr>
</tbody>
</table>

Source: Ohio Testing Data and 2007 National Assessment of Educational Progress (NAEP) Data
Table 3. Research Question Summary

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Independent Variable</th>
<th>Dependent Variable</th>
<th>Statistical Test Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>How is a district's service model related to teachers' progress with their</td>
<td>District's service model</td>
<td>Teachers' value-added math scores</td>
<td>ANOVA</td>
</tr>
<tr>
<td>gifted students in terms of value-added math scores?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are districts concerned with the ceiling effect (students topping out by hitting</td>
<td>District perceptions on the ceiling effect</td>
<td>Methods districts are using to minimize</td>
<td>Descriptive Comparison</td>
</tr>
<tr>
<td>a 99% on a test) for their gifted students, and if so, how are they attempting</td>
<td></td>
<td>the impact of the ceiling effect</td>
<td></td>
</tr>
<tr>
<td>to minimize its impact?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How are teachers' perceptions of having gifted students in their class related</td>
<td>Teachers' perceptions</td>
<td>Teachers' value-added math scores</td>
<td>ANOVA</td>
</tr>
<tr>
<td>to their value-added math scores?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How are teachers' perceptions of having gifted students in their class related</td>
<td>Teachers' perceptions</td>
<td>Teachers' OTES teacher evaluation rating</td>
<td>ANOVA</td>
</tr>
<tr>
<td>to their OTES teacher evaluation ratings?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How do gifted teachers' value-added math scores compare to their OTES teacher</td>
<td>Teachers' value-added math scores</td>
<td>Teachers' OTES teacher evaluation rating</td>
<td>T-Test</td>
</tr>
<tr>
<td>rating?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Number of Respondents, Mean Value-Added Score and Variance by Service Model Type

<table>
<thead>
<tr>
<th>Service Model</th>
<th>Number of Respondents</th>
<th>Mean Value-Added Score</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replacement Course</td>
<td>9</td>
<td>1.859</td>
<td>3.983</td>
</tr>
<tr>
<td>(Teacher of Record)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regular Classroom with HQPD</td>
<td>6</td>
<td>4.205</td>
<td>2.106</td>
</tr>
<tr>
<td>Subject Acceleration</td>
<td>5</td>
<td>2.586</td>
<td>5.523</td>
</tr>
<tr>
<td>Honors Class</td>
<td>3</td>
<td>2.637</td>
<td>9.834</td>
</tr>
</tbody>
</table>
Table 5. *Number of Respondents Concerned or Very Concerned About the Ceiling Effect and the Methods Used to Mitigate the Impact*

<table>
<thead>
<tr>
<th>Method</th>
<th>Number of Respondents</th>
<th>Method</th>
<th>Number of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tracking gifted students' value-added scores</td>
<td>N = 2, 100%</td>
<td>Tracking gifted students' value-added scores</td>
<td>N = 2, 18.2%</td>
</tr>
<tr>
<td>Math accelerating high performing students to the next grade level</td>
<td>N = 2, 100%</td>
<td>Math accelerating high performing students to the next grade level</td>
<td>N = 8, 72.7%</td>
</tr>
<tr>
<td>Pursuing professional development aimed at high performing students</td>
<td>N = 2, 100%</td>
<td>Pursuing professional development aimed at high performing students</td>
<td>N = 2, 18.2%</td>
</tr>
<tr>
<td>Other</td>
<td>N = 0, 0%</td>
<td>Other: extensions tied to above grade level standards</td>
<td>N = 1, 9.1%</td>
</tr>
<tr>
<td>Other</td>
<td>N = 0, 0%</td>
<td>Other: looking for a rigorous curriculum where student thinking, reasoning and problem solving build an understanding</td>
<td>N = 1, 9.1%</td>
</tr>
</tbody>
</table>

Table 6. *Number of Respondents, Mean Value-Added Score and Variance Based on Teacher's Perception*

<table>
<thead>
<tr>
<th>Teachers' Perceptions</th>
<th>Number of Respondents</th>
<th>Mean Value-Added Score</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel it is likely that my gifted students can show growth</td>
<td>10</td>
<td>3.139</td>
<td>6.595</td>
</tr>
<tr>
<td>I am confident that my gifted students can show growth</td>
<td>9</td>
<td>3.133</td>
<td>3.768</td>
</tr>
<tr>
<td>I am unsure if my gifted students can show growth</td>
<td>6</td>
<td>1.897</td>
<td>3.433</td>
</tr>
</tbody>
</table>
### Table 7. Number of Respondents, Mean Summative OTES Evaluation and Variance Based on Teacher's Perception

<table>
<thead>
<tr>
<th>Teachers' Perceptions</th>
<th>Number of Respondents</th>
<th>Mean Summative OTES Evaluation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel it is likely that my gifted students can show growth</td>
<td>12</td>
<td>3.5</td>
<td>0.273</td>
</tr>
<tr>
<td>I am confident that my gifted students can show growth</td>
<td>10</td>
<td>3.6</td>
<td>0.267</td>
</tr>
<tr>
<td>I am unsure if my gifted students can show growth</td>
<td>7</td>
<td>3.429</td>
<td>0.286</td>
</tr>
</tbody>
</table>

### Table 8. Number of Respondents, Mean Value-Added Score and Variance Based on Teachers' Summative OTES Evaluation Rating

<table>
<thead>
<tr>
<th>Teachers' Summative OTES Evaluation Rating</th>
<th>Number of Respondents</th>
<th>Mean Value-Added Score</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accomplished</td>
<td>13</td>
<td>3.055</td>
<td>5.438</td>
</tr>
<tr>
<td>Skilled</td>
<td>12</td>
<td>2.605</td>
<td>4.286</td>
</tr>
<tr>
<td>Developing</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Ineffective</td>
<td>0</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>