

# Can Educators Be Both Good and Successful?: The Relationship Between Socially Just (Good) and Successful Teaching

Author: Jason C. Colombino

Persistent link: <http://hdl.handle.net/2345/bc-ir:107425>

This work is posted on [eScholarship@BC](#),  
Boston College University Libraries.

---

Boston College Electronic Thesis or Dissertation, 2017

Copyright is held by the author. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<http://creativecommons.org/licenses/by-nc-nd/4.0>).

Boston College  
Lynch School of Education  
Department of Teacher Education/Special Education, Curriculum & Instruction  
Curriculum & Instruction

CAN EDUCATORS BE BOTH GOOD AND SUCCESSFUL?:  
THE RELATIONSHIP BETWEEN SOCIALLY JUST (GOOD) AND SUCCESSFUL  
TEACHING

Dissertation

by

JASON COLOMBINO

Submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy

May, 2017



### **Acknowledgements**

The completion of a dissertation is a culmination to years of hard work, persistence and thought. I would not have ever dreamed of getting to this day without the support and accountability of so many people. I would like to acknowledge them formally here.

First and foremost, to my wife and best friend, Kerry. Your unwavering support, care and love, not just in writing this paper, but in every aspect of our lives makes me a better person everyday. You mean the world to me and I would not have been able to complete this project without you. Thank you from the bottom of my heart.

To my wonderful children, Reese and Ryan, thank you for always reminding me about what is important in life and keeping me humble. I love you. To all of my parents, Mom, Dad, Dick, Mary, and Cori, thank you for showing me what it means to balance life and work and to be a good husband/father/human being. Michael, Kristin, Corrado, Doug and Holly, you are the best. Thank you for not just being a brother/sister, but true friends.

To my advisor and committee chair, Dr. David Scanlon, thank you for believing in me and not giving up on me, no matter how small the percentage seemed that I would be able to complete this project and program. I am indebted to you for your expertise and advice throughout my years at Boston College. I want to thank Dr. Richard Jackson and Dr. Eric Dearing for their critiques and scholarly opinions as readers on my dissertation committee.

To the Lynch family, thank you for providing me with financial and professional support not just as a Ph.D. student at the Lynch School, but through the Lynch Leadership Academy as well.

Thank you to the Bill & Melinda Gates Foundation for envisioning and implementing the Measures of Effective Teaching Project. Without this amazingly thought out and comprehensive database, I would not have been able to get to the findings that I did in this dissertation.

Best,

Jason

## Table of Contents

<b>Chapter 1- Introduction</b>	<b>1</b>
Problem Statement	1
Measuring Effective Teaching	2
Statement of Purpose and Research Questions	3
Overview of Methods & Contribution to the Field	5
<b>Chapter 2: Literature Review</b>	<b>6</b>
Social Justice and Socially Just Teaching	6
Opportunity to Learn as Socially Just Teaching	9
Universal Design for Learning as Socially Just Teaching	10
Connecting Socially Just Teaching, UDL and OTL	12
Theoretical and Conceptual Frameworks	13
Review of Research	14
Value-Added Models and Student Achievement	14
The Impact of Assessment and Instructional Alignment on VAM scores	17
The Impact of Specific Models on Value-Added Scores	18
Teacher Evaluation and Traditionally Underserved Populations	21
Impact on Policy and Using VAMs for High-Stakes Decisions	24
Opportunity to Learn and Student Achievement	25
Measuring and Defining the Construct of Opportunity to Learn	25
Opportunity to Learn Across Content Areas	26
Universal Design for Learning and Student Achievement	28
Social Justice and Student Achievement	29
<b>Chapter 3: Methodology</b>	<b>32</b>
Human Subjects Protection	32
Data Sources and Instruments	35
Student and Teacher Demographics	35
Student Sub-groups	39
Student Achievement	40
Classroom Observation Instruments	41
Socially Just Teaching Practices	42

Framework for Teaching (FfT)	43
Protocol for Language Arts Teaching Observation (PLATO)	45
Data Analysis Plan	45
<b>Chapter 4 - Findings</b>	<b>54</b>
Socially Just Teaching Practices	54
Findings for Research Question 1a: Socially Just Teaching and All Students	55
Findings for Research Question 1b: Socially Just Teaching and ELLs	65
Findings for Research Question 1c: Socially Just Teaching and SWDs	67
Findings for Research Question 1d: Socially Just Teaching and Students who qualify for Free/Reduced Priced Lunch	70
Findings for Research Question 1e: Socially Just Teaching and Gifted/Advanced students	72
Findings for Research Question 1f: Socially Just Teaching and Student Race	74
Findings for Research Question 2: Socially Just Teaching and Test Type	76
Findings for Research Question 3: Socially Just Teaching and Classroom Framework for Teaching Findings	78
Protocol for Language Arts Teaching Observation Findings	82
Summary of Findings	86
<b>Chapter 5 – Discussion</b>	<b>88</b>
Socially Just Teaching and Student Achievement	88
UDL and OTL and Student Achievement	92
Teacher VAMs and Student Achievement	95
Implications	95
Limitations	97
Conclusion	100
<b>References</b>	<b>102</b>

## Index of Tables and Figures

- Figure 1. *Theoretical Framework connecting socially just teaching practices and student achievement*
- Table 1. *Weights for value-added models used in the Measures of Effective Teaching Study*
- Table 2. *Frequency and percentage of participants in MET Project Videos by Subject*
- Table 3. *Frequency and percentage of participants in MET Project Videos: by Grade*
- Table 4. *Teacher Descriptive Statistics for Subsample*
- Table 5. *Student Descriptive Statistics for Subsample*
- Table 6. *FtF Components included in MET Video Scoring and Alignment to UDL Principles*
- Table 7. *Summary of Regression Models*
- Table 8. *Item Mean Scores and Overall Average on Teaching for Social Justice Observation Scale (TSJOS)*
- Table 9. *Regression #1: Predicting z-score for section average 2011 ELA state exam from mean TSJOS score.*
- Figure 2: *Student achievement predicted by socially just teaching practices*
- Table 10. *Regression #2: Predicting student rank-based z-score from 2011 state ELA exam from mean TSJOS score.*
- Table 11. *Regression #3: Predicting student achievement for all students from location, teacher, student and socially just teaching variables.*
- Table 12. *Regression #4: Predicting student achievement for 4th and 5th grade classrooms from socially just teaching scores.*
- Table 13. *Regression #5: Predicting student achievement (SAT9) for 4th and 5th grade classrooms from socially just teaching scores.*
- Table 14. *Regression #6: Predicting student achievement for ELLs from socially just teaching scores.*
- Table 15. *Regression #7: Predicting student achievement for SWDs from socially just teaching scores.*
- Table 16. *Regression #8: Predicting student achievement for students who qualify for free/reduced price lunch from socially just teaching scores.*
- Table 17. *Regression #9: Predicting student achievement for gifted/advanced students from socially just teaching scores.*
- Table 18. *Regression #10: Predicting student achievement for Black, Hispanic or American Indian students from socially just teaching scores.*
- Table 19. *Regression #11: Predicting student z-score scaled scores on SAT9 from mean TSJOS score.*

Table 20. *Regression #12: Predicting student achievement from average ELA FfT scores and mean TSJOS scores.*

Table 21. *Regression #13: Predicting student achievement from location and teacher variables, average ELA FfT scores and mean TSJOS scores.*

Table 22. *Regression #14: Predicting student achievement from average PLATO scores and mean TSJOS scores.*

Table 23. *Regression #15: Predicting student achievement from location and teacher variables, average PLATO scores and mean TSJOS scores.*



**Abstract**

Can educators be both good and successful?:

The relationship between socially just (good) and successful teaching

Jason C. Colombino

Dissertation Advisor: Dr. David Scanlon

There is limited research on the relationship between socially just teaching practices and student achievement. While successful teaching is often defined through test scores, good teaching encompasses the moral elements of teaching (Fenstermacher & Richardson, 2005). This study, building on the work of Mitescu, Cochran-Smith, Pedulla, Cannady, and Jong (2011), is a secondary analysis examining the relationship between socially just teaching practices and student achievement. A subsample of 4th and 5th grade English/language arts (ELA) teachers (n=107) and students (n=2587) was taken from the Measures of Effective Teaching Longitudinal Database (Bill and Melinda Gates Foundation, 2013). Classroom videos were coded using the Teaching for Social Justice Observation Scale (TSJOS) of the Reformed Teaching Observation Protocol-Plus (RTOP+) (Mitescu et al.) to measure socially just teaching practices.

Unadjusted linear regression analyses indicated a positive significant correlation between teachers' mean TSJOS score and the class averages on standardized state ELA exams and the class average on an assessment of higher-order thinking skills. This relationship was also found when the same analysis was conducted on 4th grade classrooms as well as 5th grade classrooms. A hierarchical multiple linear regression found a positive significant relationship between TSJOS scores and student achievement after accounting for location, teacher, and student predictor variables. The relationship between socially just teaching practices and student achievement for subgroups of students is discussed.

The study analyzed the significance and magnitude of the relationship between socially just teaching practices after two widely used classroom observation protocols, the Framework for Teaching (FfT) and the Protocol for Language Arts Observation Scale (PLATO), were entered into the model. Teacher mean TSJOS scores were found to explain a significant and unique proportion of the variation in state assessment scores after accounting for average FfT ELA observation scores and teacher average PLATO observation scores, separately.

This study adds to the literature on the connection between socially just teaching practices and student achievement, in that it provides compelling evidence that socially just teaching practices are not only related to the good, or moral, side of teaching, but also have a positive and significant relationship with increased student achievement for all students.



## Chapter 1 - Introduction

### Problem Statement

Throughout its history, the United States education system has grappled with how to commonly define what is meant by “excellent teaching,” with a high degree of reliability and validity. Value-added models (VAMs) are one popular approach to quantify the contribution a particular teacher has on his or her students (see Braun [2005] for an overview of value-added-models). Value added models are not without problems, including inconsistency and that teachers’ scores are affected by the students that they are assigned, and VAMs cannot differentiate the influence of other variables that impact student achievement (Darling-Hammond et al., 2012). Value-added models, however, that utilize a combination of observation scores, student achievement scores and student perception surveys, have shown that teachers’ impact on student achievement can be measured and identified (Bill and Melinda Gates Foundation, 2013).

An underlying assumption in the teacher evaluation process is that effective teaching is equated to improving student outcome on standardized instruments that measure academic achievement. While some argue it is appropriate to define successful teaching primarily in terms of academic achievement (Hanushek, 2002), others (e.g., Fenstermacher & Richardson, 2005) argue that teaching can be split up into the intentions of the teacher, or the task, and the intended outcomes, or achievement. A teacher, for example, can meet the intended outcomes of raising achievement scores, but have immoral teaching methods. Students deserve not only a teacher who is successful in

meeting academic goals, but also a good teacher who teaches following good and moral intentions (Fenstermacher, & Richardson, 2005). Another important idea from Fenstermacher and Richardson is there are a number of variables that affect student achievement that are out of the control of teachers. Examples of such variables include whether students qualify for free or reduced price lunch, their disability or ELL status and students' previous year test scores (Mihaly, McCaffrey, Staiger, and Lockwood, 2013). In this study, socially just teaching encompasses both good and successful teaching, as described by Fenstermacher and Richardson.

There is evidence to suggest that good teaching has a positive effect on student achievement, counter to the argument that socially just teaching can be a barrier to, rather than bolster, student achievement (Stotsky, 2009). A significant correlation between a measure of socially just teaching and student achievement was found in a study of elementary math classrooms (Mitescu, Cochran-Smith, Pedulla, Cannady, & Jong, 2011), for example. In a review of the literature on social justice in teacher education, however, Grant and Agosto (2008) found that the term "social justice" is typically not fully defined in research articles, and that there is limited research on assessing social justice teacher preparation programs. This study analyses the relationship between socially just teaching practices and student achievement outcomes.

### **Measuring Effective Teaching**

Teacher evaluation systems attempt to identify effective teaching. Recently, evaluation systems have put an emphasis on including multiple measures to quantify

teacher effectiveness. For example, a measure of teacher impact on student achievement is a part of many state valuation systems. Student achievement is often defined in terms of standardized assessments, including the Partnership for Assessment of College and Career Readiness (PARCC). Student feedback on teacher's pedagogy and classroom management and classroom observations are two other elements commonly found in teacher evaluation systems.

The Measures of Effective Teaching (MET) database (Bill and Melinda Gates Foundation, 2010a) has been created to identify effective teaching practices linked to student achievement. The MET project found that classroom observations have been found to measure part of teachers' effectiveness, in terms of explaining a portion of the variance in student test scores; multiple observations by multiple observers providing a higher explanatory power relative to one observer. A recent study (Grossman et. al., 2014) found that the value added scores for teachers changed depending on the type of assessment used. In this study, one particular observation protocol, the PLATO (Protocol for Language Arts Teaching Observation), was related more strongly to an assessment aimed at assessing higher-order thinking compared to traditional standardized state assessments. Teacher evaluation systems are a crucial piece of the recent reform effort for identifying good teaching and evaluation systems that make connections between teaching and student achievement.

### **Statement of Purpose and Research Questions**

The intent of this research study is to outline the research topic and describe the relationship and impact between socially just, or good, teaching and successful (as assessed by student achievement) teaching. The study addresses the following research questions, based on a review of current literature in the field:

1: What is the relationship between teacher Social Justice classroom observation scores and student achievement for...

- a. All students?
- b. English Language Learners (ELL) (and non-ELL students)
- c. Students with Disabilities (SWD) (and non-SWD)
- d. Students receiving free or reduced price lunch? (and students who do not receive free/reduced price lunch)
- e. Gifted/Advanced students (and non-gifted/advanced students)
- f. Students who are Black, Hispanic, or American Indian (and students who are not Black, Hispanic or American Indian).

2: What is the impact of the type of student achievement measure (standardized state ELA exams versus the Stanford Achievement Test 9/Open Ended Reading assessment) on the relationship between teacher social justice classroom observation scores and student achievement?

3: What, if any additional explanation of variance in student achievement scores can be explained by teacher social justice classroom observation scores above and beyond the variance in scores explained by (i.) Average Framework for Teaching (FfT2) ELA

observation scores and (ii.) Average Protocol for Language Arts Teaching Observation (PLATO) scores.

### **Overview of Methods & Contribution to the Field**

There is limited research analyzing the relationship between teaching practices related to social justice and student achievement (Mitescu et al., 2011). This study adds to the current bodies of research on teacher evaluation and the relationship between socially just teaching and student achievement. The relationship between socially just teaching practices and student achievement will be assessed through linear and hierarchical multiple linear regression analysis. Evidence of a relationship will be a significant, positive relationship between socially just teaching practices and student achievement. The impact of a relationship, then, will be assessed by the change in the proportion of variance, measured by the  $R^2$  change statistic, when the measurement of socially just teaching practices is added to the analysis as the last variable in the last block of variables for each analysis. This study makes a unique contribution by drawing from a sample that includes a random assignment of class lists to teachers, thus allowing ruling out some variation in student achievement measures due to other teacher and student level variables; in this way it further explains the relationship between socially just teaching practices and student achievement.

## Chapter 2 - Review of the Literature

Socially just teaching matters. While schools should have a mission that focuses on increasing student achievement, which is one conceptualization of successful teaching, that should be one of the many outcomes of socially just teaching. Moral and ethical, or good, teaching is another piece of socially just teaching. Therefore, socially just teaching includes both good and successful teaching, as defined by Fenstermacher and Richardson, (2005). The following review of the literature frames the three main components of the conceptual framework for this study (social justice, UDL and OTL) as socially just teaching and includes relevant research on each topic.

### **Social Justice and Socially Just Teaching**

As Collopy, Bowman and Taylor (2012) argue “The educational achievement gap is a critical social justice issue” (p. 4). The English philosopher John Stuart Mill is credited with a definition of social justice that influenced current education theorists:

Society should treat all equally well who have *deserved* equally well of it, that is, who have deserved equally well absolutely. This is the highest abstract standard of social and distributive justice; towards which all institutions, and efforts of all virtuous citizens, should be made in the utmost degree to converge (as cited in Novack, 2000, p. 12).

Stuart Mill’s definition of social justice includes both socially just people, and institutions and organizations. While he emphasizes equity, the idea of equity rather than equality is referenced in modern definitions of social justice (Goodlad [2002] and Barry [2005] as cited in Novack, 2000). Rawls (1971 as cited in Templeton, 2011) includes the notion of fairness as a cornerstone of the definition of social justice. The deliberate intent of



teachers to critique knowledge and meaning making is also included in modern definitions of social justice:

“Learning to teach that is premised on a stance for social justice recognizes the importance of social justice pedagogy. This social justice pedagogy refers to a deliberate attempt to construct authentic conditions through which educators and students can think critically about what stands as knowledge, how knowledge is produced, and how knowledge is transformed by a particular relationship between the self, others, and the larger world’ (Goroux, 1992 as cited in Grant and Agosto, 2008).

An important distinction can also be made between two modern era theorists, Freire (1993) and Gramsci (1971 as cited in Maranto and Ritter, 2014), in the search for a definition of social justice, in terms of the goal or desired outcome of socially just teaching practices. Freire argues for the oppressed to overthrow the current system of privilege. Gramsci differs from Freire, as he advocates for the oppressed to learn to live within the culture, norms and language of the system of privilege in order to allow the oppressed to have the choice and power to make changes to the current system as a means to make it a more just and fair society. There is an argument that while the educational research community, including the American Education Research Association, are more aligned with Freire’s goals for social justice, many advocacy groups, including the Black Alliance for Education Options and the Democrats for Education Reform follow Gramsci’s philosophy (Maranto and Ritter, 2014). No matter what particular definition of social justice is utilized, however, this study assumes that

increased student achievement is one goal of social justice as improved academic achievement will open doors and opportunities for students.

Remedying the educational achievement gap is one goal of socially just teaching, as improved achievement, will, in theory, provide students with more equitable access to the knowledge and skills held by the privileged. Socially just teaching aims to move beyond achievement, however, by providing students and teachers with tools to reflect and critique knowledge as well as to look at systems of privilege and oppression. In *The Dreamkeepers*, for example, Ladson-Billings (2009) describes culturally relevant teaching by eight teachers who held high expectations for African American students, and emphasizes that it is the way that those teacher taught that made all the difference for students.

To recap, socially just teaching is teaching that has an aim at not only eliminating the achievement gap, but eliminating the achievement gap by providing each student with an equitable “leveling culture of achievement that extends to all of its members and a strong sense of group membership, where the expectation that everyone achieves is explicit and is regularly communicated in public and group settings” (Perry, Steele and Hilliard, 2003, p. 107). The end result of socially just teaching is “meeting everyone’s basic needs and fulfilling everyone’s potential to live productive and empowered lives as participating citizens of our global community” (Wade, 2007, p. 5).

For the purposes of this dissertation study, social justice at its core is about fairness and equity for all students. Socially just teaching is defined through the 14 items

on the Teaching for Social Justice Observation Scale (TSJOS) of the Reformed Teaching Observation Protocol-Plus (RTOP+) (Mitescu et al., 2011) that will be used to measure socially just teaching practices.

### **Opportunity to Learn as Socially Just Teaching**

Providing students with an Opportunity to Learn (OTL) is one of the most important issues in US education law, policy and practice. Until recently, OTL has been commonly equated with scores on statewide assessments (MCAS in Massachusetts, for example). A more nuanced definition of OTL (Moss, Pullin, Gee, Haertel and Young, 2008) conceptualizes OTL as an intersection between the student and his or her learning environment. OTL takes the middle road between the “historical highway” view of disabilities as residing in an individual and the more “back road” post-modernist interpretation of disabilities as being a manifestation of the particular culture and environment in which an individual resides (Moss et al, 2008). OTL, then, is one conceptual framework that aligns with the goals of social justice and socially just teaching practices.

In the proposed study, students’ OTL will be measured through both the traditional, test-score driven, definition of OTL as well as an assessment of students’ OTL through the re-conceptualized notion of OTL as defined by Moss et al. More specifically, teacher content knowledge will inform the estimate of teacher effectiveness in addition to traditional standardized common assessments. Hence, OTL in the present study can be analyzed relative to curriculum content, resources, teaching practices, and

the links among all three and the individual learner (Pullin and Haertel, 2008). Utilizing an OTL lens will uncover the current state of socially just teaching practices and issues of equity more clearly.

### **Universal Design for Learning (UDL) as Socially Just Teaching**

Universal Design for Learning (UDL) is a powerful, integrated framework that shifts the problem of learning from that of a problem with the child to a problem with the curriculum (Hitchcock, Meyer, Rose, & Jackson, 2002). UDL has its origins in the work of Mace, who pioneered architectural work on designing buildings using universal design to make buildings more accessible to a wider range of people (Jimenez, Graf, & Rose, 2007). Universal Design for Learning has been defined in the Higher Education

Opportunity Act of 2008 (section 103(a)(24)):

The term “UNIVERSAL DESIGN FOR LEARNING means a scientifically valid framework for guiding educational practice that:

(A) provides flexibility in the ways information is presented, in the ways students respond or demonstrate knowledge and skills, and in the ways students are engaged; and

(B) reduces barriers in instruction, provides appropriate accommodations, supports, and challenges, and maintains high achievement expectations for all students, including students with disabilities and students who are limited English proficient.”

There are three guiding principles to UDL: provide multiple means of (1. representation, (2. action and expression, and (3. engagement (CAST, 2011).

Universal Design for Learning has the potential to integrate the current dualism of general and special education (Kauffman, 2016; Poplin, 2005), or in the postsecondary context, general education and students receiving accommodations. UDL provides a more nuanced and deeper meaning of inclusion than how inclusion has been defined in the past (Kavale and Forness, 2000). For example, in UDL inclusion means more than the placement of students with disabilities in a class with students without disabilities, it also means students with disabilities (and their peers) receiving individualized instruction that maximizes their opportunity to learn. At its core, UDL is an inherently socially just framework for developing curriculum, as its goal is to provide flexible and accessible curriculum, instruction, and assessment for all students.

Universal Design for Learning attempts to integrate and synthesize research findings on effective practices from various disciplines, most notably neuroscience, cognitive psychology and education. Going beyond the medical model of disabilities, which emphasizes neurological causes for disabilities (Poplin, 1988), UDL places emphasis on the interaction between children and their environment, rather than focusing on disabilities as being something only internal to a student (CAST, 2007). Analyzing curriculum and instructional practices in postsecondary classrooms in terms of the three principles of UDL (multiple means of representation, expression, and engagement) (Rose & Meyer, 2002) can help filter the data to identify relative strengths and weakness of the curriculum in terms of accessibility for student learning. Analyzing the interaction between the curriculum and students with higher functioning forms of Autism, for

example, provides an example of examining students “at the margin” (Rose & Meyer, n.d.) where many innovative curriculum adaptations and improvements are identified that ultimately benefit a large portion of students. In this study, the three main tenants of UDL are used as an observation tool to measure teachers’ impact on student learning. An example of schools applying UDL principles to teacher observation and evaluation is the Bartholomew Consolidated School District (BCSD), a school district serving over 12,000 students. BCSD has taken steps to implement a teacher evaluation system utilizing UDL principles. In fact, 50% of every teacher’s evaluation rating is dependent on a UDL rubric developed by the district (Ganley and Ralabate, 2013).

### **Connecting Social Justice, Opportunity to Learn and Universal Design for Learning**

Taken together, OTL and UDL provide a particular lens, or unique stance, to inquire into what is happening at the intersections of student learning, socially just instructional practices and policy. The integration of these frameworks allows for robust connections and explanations between good and successful teaching. Findings from this study have the potential to identify recommendations for policy-makers and practitioners on how to improve access, equity, and OTL for all students, particularly traditionally underserved and at-risk students. The framework allow for the valid critique of the accessibility of curriculum for students with a plethora of learning profiles.

To summarize, OTL and UDL can, and have been, described as fitting into the overarching objective of social justice for all learners. For example, the theme of the 2017 UDL Annual Symposium is “UDL for Social Justice,” indicating that the research

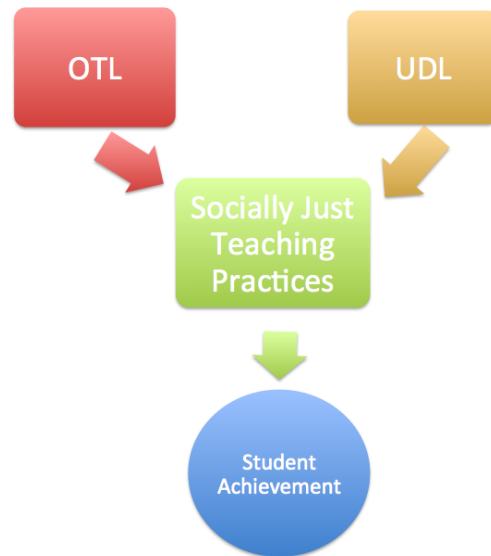
field has embraced the notion that UDL is one framework that can be utilized to make progress toward equity and social justice for all students. The tenets of OTL and UDL, taken together, align with socially just teaching practices, as defined by the 14 items on the TSJOS classroom observation tool.

### **Theoretical and Conceptual Frameworks**

The theoretical framework acts as a filter for interpreting data and “is yet another way in which a researcher makes his or her findings intelligible to an academic audience and open to scrutiny” (Kilbourn, 2006, p. 545). In the present study the theoretical frameworks of Opportunity to Learn (OTL) and Universal Design for Learning (UDL) provide distinct analytic lenses to analyze and critique the opportunity to learn and provide a social justice perspective to analyzing the impact of teaching on student achievement (see Figure 1). Opportunity to Learn provides clarity on the definition of socially just teaching, while UDL offers insight to principles of socially just teaching in terms of curricular, instructional, and assessment design. The theoretical framework for this proposal uses the two constructs of OTL and UDL to define socially just teaching practices. The overarching objective of the study is to describe the relationship and impact of socially just teaching practices on student achievement.

**Figure 1:**

*Theoretical Framework connecting socially just teaching practices and student achievement*



### Review of Research

This particular literature review can be best described as a traditional narrative review. The criteria provided by Boote and Beile (2005) was used as an evaluative tool during the writing process to enhance the quality of the literature review. The literature review analyzes what the research says about teacher evaluation and student achievement. Topics for the literature review are as follows: Value-Added Models, Opportunity to Learn and Student Achievement, Universal Design for Learning and Student Achievement, and Social Justice and Student Achievement.

**Value-Added Models and Student Achievement.** The review of literature on VAMs and student achievement is organized into four topics of inquiry, all of which serve a specific purpose for this proposed dissertation:

- 1. The impact of assessment and instructional alignment on VAM scores*
- 2. The impact of varying models on value-added scores*



*3. Teacher Evaluation and Traditionally Underserved Populations*

*4. Impact on policy and using VAMs for high-stakes decisions*

Value-added models (VAMs) are a statistical approach to quantify the contribution of an educator's teaching to student achievement. As Braun (2005) notes, a major concern with associating causality with VAMs is the typically incorrect assumption that students are randomly assigned to classes and teachers as well. In fact, research suggests that students typically are not randomly assigned and administrators, teachers and parents play a role in classroom placement (Paufler & Amrein-Beardsley, 2014). While inferring causality from most VAMs is not a valid way to determine a teacher's effectiveness or impact on student achievement, VAMs have shown promise in estimating the contribution that teachers make to their students' progress. There are a number of different VAMs that take student, school, and teacher characteristics into account in attempt to explain some of the variability in student scores, to minimize the error associated with the student and teacher placements not being random. In essence, many VAMs run a multiple regression with a multitude of student, teacher and school-level variables as inputs to attain an estimated output score, in most cases student scores on standardized assessments.

The Measures of Effective Teaching (MET) project addressed the major concern of random assignment associated with VAMs. In the MET study, approximately 3,000 teachers from six different districts (the Charlotte-Mecklenburg Schools, Dallas Independent Schools, Denver Public Schools, Hillsborough Public Schools, Memphis

Public Schools and Pittsburgh Public Schools) volunteered to be part of the project. Class lists were randomly assigned to teachers and student outcomes were predicted using factors like a teacher's previous record of testing results, observations and student feedback. The predictions were then compared with student achievement results.

The major findings from the MET project included that effective teachers could be identified. The project used four different models, incorporating different percentages, or weights, to different measurements. The predictability and variability of scores differed based on the weighting model being used. In terms of reliability, the MET project found that different observational setups impact the reliability of classroom observations, based both on number of classroom visits and who is observing (administrator or peer) (Bill and Melinda Gates Foundation, 2013).

The MET study reduced the major barrier of random assignment in teacher evaluation with convincing results that good teaching can be estimated and measured with validity and reliability. Teacher evaluation is a major research topic and the intent of this literature review is to narrow the focus to topics related to socially just teaching through a UDL/OTL perspective. The literature review first focuses on research related to the impact of student assessments and instructional alignment on value-added models ties to socially just teaching as assessments and instruction are two ways to analyze OTL. Second, the impact of the specific VAM to be used to evaluate teaching is reviewed. If socially just teaching practices are not captured in VAM then the field is limited in what can be said about the relationship between socially just teaching and student achievement

and growth. Third, the state of current research on VAM for traditionally underserved student populations is addressed. Fourth, the impact of VAM on policy and high-stakes hiring/firing decisions is considered, as it is central to this topic. If a link between socially just teaching and student achievement can be identified, and if teachers can be evaluated accurately on the extent and quality of their socially just teaching practices, then policy can be put in place to not only advocate for socially just teaching practices, but to hold districts, schools and teachers accountable for these practices of equity.

***The Impact of Assessment and Instructional Alignment on Value-Added Models.*** While VAMs typically utilize a standardized assessment as the desired outcome variable of interest, the specific type of assessment can impact teacher effectiveness ratings. Three different reading achievement tests, for example, were found to have only moderate correlations with teacher value-added scores (Papay, 2011). The six state tests used in the MET study differed significantly in their correlations to VAM scores, and the correlations to instructional quality were generally low (Polikoff, 2014). One study (Grossman, Cohen, Ronfeldt & Brown, 2014) assessed the relationship between a specific classroom observation protocol used in the MET study, the Protocol for Language Arts Teaching Observation (PLATO), and student achievement on state assessments and an assessment of higher order thinking (the Stanford Achievement Test, SAT-9). The PLATO correlated more with the SAT-9 than the state achievement tests. The SAT-9 was also found to be more sensitive to the PLATO index of Cognitive and Disciplinary Demand than the state test. These findings, collectively, suggest that teacher

observations are sensitive to the type of test used to measure student achievement in value-added models, and there are differences in instructional sensitivity from state to state.

While there is emerging research suggesting that value-added scores are sensitive to the specific test used to measure student achievement, little research has been completed analyzing the impact of teachers' instructional alignment to VAM scores. Polikoff and Porter (2014) used the MET database and found that teachers' instructional alignment had a weak correlation to value-added scores. The most plausible hypothesis for the lack of association is that state tests do not accurately measure day-to-day teaching and learning that occurs in the classroom. This study assesses the differences in regression analyses that use a measure of socially just teaching practices to predict student achievement for two different measures of student achievement.

*The Impact of Specific Models on Value-Added Scores.* Teacher VAM scores depend not only on the specific outcome measure of student achievement, but also on the input variables and weights associated with those variables. The MET project (Bill and Melinda Gates Foundation, 2013) investigated the correlations with state and higher-order tests as well as the reliability of four models. Each of the four models placed different weights on student test gains, student surveys, and classroom observations (see Table 1). Model 1 consisted of weights that maximized the correlation with student achievement.

**Table 1**

*Weights for value-added models used in the Measures of Effective Teaching Study*

<b>Model</b>	<b>Student Achievement</b>	<b>Student Surveys</b>	<b>Classroom Observations</b>
1	50%	25%	25%
2	81%	17%	2%
3	33%	33%	33%
4	25%	25%	50%

(Bill and Melinda Gates Foundation, 2013, p. 11)

Results indicate that the different measurements used in the MET project had overlap in terms of explaining some of the teacher effectiveness (Bill and Melinda Gates Foundation, 2013), which aligns with findings that VAM scores of teachers are correlated to their content knowledge, observations and characteristics of students they teach (Hill, Kapitula & Umland, 2011). Which VAM model that is used, however, has been shown to result in significant differences for teacher effectiveness ratings at the high school level (Goldhaber, Goldschmidt, and Tseng, 2013) and in terms of how missing data is treated (Karl, Yang and Lohr, 2013). In the MET 2013 report, the correlation to state tests decreased as the weight placed on student test gains decreased, but the correlations to higher-order tests were relatively flat (ranging from 0.29 to 0.34). The reliability had an inverse relationship as the reliability increased as the weight placed on student test gains decreased. These findings, as noted in the MET technical report by Mihaly, McCaffrey, Stiager, and Lockwood (2013), indicate there are costs and benefits to different weighting models when calculating overall teacher value-added scores. The findings from the MET project demonstrate the sensitivity of VAM scores to different

weighting models, and that states, districts and schools should look at decisions around these issues based on the priorities and local context of the evaluation system being used.

The reliability of the inputs of VAM models also deserves consideration. Classroom observations, for example, have been shown to have different levels of variability based both on the duration of the observation as well as on who is conducting the observation, and using an increased number of observers for any given number of observations increases reliability (Bill and Melinda Gates Foundation, 2013).

Observations that occur for the first thirty minutes of a class are predictive of observation scores for a full lesson. In addition, inter-rater reliability for classroom observations is higher when a smaller number of related observation items are scored compared to using a full observation protocol (Joe, McClellan, Holtzman, 2014).

In terms of student test scores, VAM scores that use test results from multiple subjects do a better job of identifying a teacher's effectiveness in the future compared to models that use only one test subject score (Lefgren & Sims, 2012). This finding is applicable to elementary school settings, where teachers typically are responsible for teaching more than one subject. Combining these test scores to estimate a teacher's effectiveness is one way to improve predictability of VAM scores, especially in light of the finding that the multiple subject models are even more robust when there is a limited number of years of student data available (Lefgren & Sims, 2012). At the high school and college level, multivariable value-added assessment models have been created that include both binary outcomes, like graduating with a certain major, and continuous

variables, including test scores, to give a more holistic view of a teacher's impact and effectiveness (Broatch & Lohr, 2012).

The current project utilized current research on VAM models to develop regression models to explore the relationship between socially just teaching practices and student achievement. The conceptual framework of integrating an OTL and UDL approach to define socially just teaching practices, which in turn will be used in developing the models used in the regression analyses. These steps will add to the current body of research around describing the relationship between socially just teaching practices and student achievement.

***Teacher Evaluation and Traditionally Underserved Populations.*** Social justice and OTL are two main driving concepts for this research, as both have the potential to add to the current body of research around teacher evaluation systems for traditionally underserved populations. If we can improve our ability to identify effective teachers for Students with Disabilities (SWD), English Language Learners (ELLs) and others who are traditionally underserved then there may be increased opportunities to identify effective teaching practices and improve achievement for these students.

There are a number of issues around both VAM scores and classroom observations for SWD and ELL students. Students with disabilities and ELL students typically have lower scores and extreme scores tend to have higher variability. Also, observation protocols typically fail to include specific instructional strategies found to be effective for teaching SWDs and ELLs (Jones, Buzick, & Turkan, 2013).

Recommendations to improve teacher evaluation for SWDs and ELLS include validating observation protocols for these sub-groups (Jones & Brownell, 2014; Jones, et al., 2012) and to consider the variability of learning profiles and achievement within the sub-groups (Jones et al., 2013). Teachers that are effective teachers of ELLs generally are also effective with non-ELL students (Loeb, Soland & Fox, 2014). The findings also indicate, however, that some teachers are relatively more effective for ELL students than non-ELL students.

Current research indicates that teacher evaluation for SWDs and ELLs is an area in need of further research. This proposed study will analyze student achievement scores for these sub-groups based on a Social Justice observation score, derived from the Teaching for Social Justice Observation Scale (TSJOS). Achievement scores will also be analyzed as a way of assessing the effect of socially just teaching practices on both SWD and ELL achievement.

The literature also speaks to the connection between social justice and educational outcomes for SWDs and ELLs. Social justice has been proposed as an umbrella construct to analyze and improve outcomes for SWDs in general (Kauffman, Anastasiou, Badar, Travers, & Wiley, 2016), as well as in specific contexts like charter schools (Shealey, Sparks, and Thomas, 2012); it is also described as a goal that can be reached through specific inclusive practices (Baglieri & Knopf, 2004). It has further been argued that in order to increase equity for SWD the focus needs to be on instruction, and not solely on inclusion, the same is true to increase student achievement (Kauffman & Badar, 2014).



On the other hand, Castro-Villarreal and Nichols (2016) argue that standardized testing required for accountability has had a negative impact on social justice and opportunities for students with disabilities. Instructional structures and interventions like Response to Intervention (RTI) has been described as having socially just outcomes, in that RTI can improve opportunities for SWD and students from traditionally marginalized races (Artiles, Bal, and Thorius, 2010).

The concept of social justice, through leveraging school leaders' point of view, and actions regarding, ELL, can increase opportunity and equity (Brooks, Adams, Morita-Mullaney, 2010; Theoharis & O'Toole, 2011). Variables like location and placement of ELL students has been argued to be evidence of, or lack of, social justice (Skinner, 2012). One means to improve social justice goals for students from different cultures is to teach bicultural teacher candidates the perspective of social justice, as this increases opportunities and equity (Weisman, Flores, and Valenciana, 2008). The connection between ELL and SWD achievement as not only a social justice issue, but an issue that can be improved through a social justice framework, supports the analysis that follows in this study to analyze the relationship between socially just teaching practices and student achievement for traditionally underserved populations, including SWDs and ELLs.

This study adds to the current knowledge base on teacher evaluation relative to traditionally underserved populations, by analyzing student achievement scores based on classroom observations measuring teacher socially just teaching practices and

achievement tests for SWDs, ELLs, students of color, students eligible for free or reduced price lunch, and gifted/talented students, all from an OTL/UDL perspective. The underlying premise is if instruction and assessment are more flexible and accessible in ways that reflect socially just teaching to a wider range of learners, then all learners, including the sub-groups, will benefit more compared to a relatively more static model of instruction and assessment.

***Impact on Policy and Using VAMs for High-Stakes Decisions.*** The use of VAM scores to make hiring and firing decisions is a practice that has been employed at least partially by a number of states (Winters & Cowen, 2013). There are a number of policy implications to the use of VAM scores for teachers, schools, districts and states. Winters and Cowen (2013) concluded that teachers that would have been fired based on VAM scores alone had students with lower gains on achievement tests compared to teachers who would have been retained using only VAM scores for firing decisions. In addition, research shows that at least three years of student data need to be used in calculating teacher VAM scores in order to keep overall error rates at or below 10% (Schochet & Chiang, 2013), which still seems like a relatively large error rate or percentage considering the high-stakes nature of the decision to lay a teacher off. VAM scores have a positive, but weak correlation (0.276 in math and 0.168 in reading), to classroom observations conducted by principals (Harris, Ingle & Rutledge, 2014).

Taken together, the policy implications of using VAM scores to inform hiring and firing decisions continues to be an area in need of further research. This study includes

policy implications, particularly in regards to the impact of using an OTL/UDL approach to teacher evaluation, and/or the use of observational tools measuring socially just teaching practices to provide teachers with feedback on areas of strength and improvement.

**Opportunity to Learn and Student Achievement.** A literature review was conducted in May, 2016 using the Eric Research Complete database. The Boolean search terms “opportunity to learn” AND “(student achievement or academic achievement)” were used and the search was limited to peer reviewed publications. The search resulted in 95 publications. After reviewing the article titles and abstracts, the author identified 24 publications that were relevant to this proposal. The decision to include a publication was based on the fact that the publication directly addressed either (1. measuring or defining the construct of OTL, or (2. analyzing the relationship between OTL and student achievement.

***Measuring and Defining the Construct of Opportunity to Learn.*** As previously stated, OTL is defined for the purposes of this proposal as the interaction or intersection between a student and his/her learning environment (Mosset al., 2008). OTL has traditionally been defined in terms of instructional time (Jaafar, 2006). OTL has also been defined using factor analysis to identify variables that impact students with and without disabilities differently (Cawthon, Beretvas, Kaye and Lockhart, 2012). There was a differential effect on student reading achievement for classroom activities reported by students and student constructed projects, and a differential effect on student math

achievement for using a calculator between students with and without disabilities. This highlights the possibility of identifying multiple factors to define OTL and ways to measure the differential impact of OTL on student achievement for students with and without disabilities. The inclusion of specific variables, including academic language instruction, has been advocated in defining OTL for English Language Learners (Aguirre-Munoz and Amabisca, 2010).

OTL has also been defined through analysis of classroom artifacts as a way to measure OTL that differs from traditional measurement based on instructional time (Jaafar, 2006). The notion of the intended, planned and enacted curriculum (Kurz, Elliott, Wehby, and Smithson, 2010) and assessment (Kurz, Talapatra, and Roach, 2012) is yet another frame by which OTL has been defined and measured in the field.

While the goals of OTL center around providing equitable educational opportunities for all students, critiques of OTL hinge on measurement difficulties (Elliott, 2015) and not including variables outside of school that impact student achievement (Dougherty, 1995). The critiques and advances around the definition of OTL align with the needs for additional research on OTL as a construct and its impact on student achievement.

***Opportunity to Learn Across the Content Areas.*** There are a number of studies that assess the relationship between OTL and student achievement in specific content areas. In terms of mathematics, the organization of high school mathematics curriculum has been found to be related to student achievement, with students choosing to take an

integrated course scoring significantly higher than students who enrolled in a traditional geometry course (Tarr, Grouws, Chavez, and Soria, 2013). In this study, OTL was included as a teacher-level variable and found to have significant predictive power in terms of explaining variability of student achievement. Socioeconomic status and friends has also been used to define OTL and found to have a direct impact on rural students' math achievement in high school (Reeves, 2012). Teacher instructional moves have been found to be a significant predictor of math achievement in both South African and Botswanan schools (Reeves, Carnoy, and Addy, 2013). Textbook content (Tornroos, 2005) is highly correlated with student achievement, and the middle school courses students take have an impact on their high school math achievement (Wang and Goldschmidt, 2003). Finally, Albano and Rodriguez (2013) found a positive relationship between preservice teachers' perception of OTL and student achievement on the Math Content Knowledge (MCK) assessment.

Research has also shown a relationship between OTL and student math achievement for ELL students, finding that tracking ELL students predicted student achievement better than student English language proficiency (Callahan, 2005). Additionally, among students from low-income families, an increase in OTL elementary mathematics was correlated with increased math achievement scores (Wang, 2009); OTL, language proficiency and immigrant status were found to be correlated with math achievement (Wang and Goldsmchidt, 1999).

Research in content areas besides mathematics shows a similar positive

relationship between OTL and student achievement. OTL has also been connected with increased student achievement in science (Mo, Singh, and Chang, 2013; Wang, 1998) and US History (Heafner and Fitchett, 2015). Research involving student and teacher perceptions of OTL has also been linked to student achievement (Cooper and Liou, 2007; Scherff and Piazza, 2008; Wiley and Yoon, 1995). There are similar findings for specific academic skills. In reading, for example, students' OTL fiction and nonfiction reading skills were found to be correlated to their achievement (LaFontaine, Baye, Vieluf, and Monseur, 2015), and reduced class sizes in middle school was found to be related to student writing achievement (Tienken and Achilles, 2009). {Jason- it didn't quite work to call reading & writing content-area}

Taken together, the cumulative research on OTL and student achievement has two main takeaways relative to the present study. The general positive relationship between OTL and student achievement, across grade levels and content areas, has been established in the research literature. A second takeaway is there is limited to no research on OTL and student achievement that utilizes a random assignment sampling methodology. This study adds to existing literature on OTL and student achievement by analyzing social justice as an OTL factor through a secondary analysis of a data set that includes a random assignment of student class rosters to teachers.

**Universal Design for Learning and Student Achievement.** A literature review was conducted in May, 2016 using the Eric Research Complete database. The boolean search terms “Universal Design for Learning or UDL” AND (student achievement or

academic achievement or academic performance) were used and the search was limited to peer reviewed publications. The search resulted in four identified publications. After reviewing the article titles and abstracts, the author identified publications that were relevant to this proposal. The decision to include a publication was based on the fact that the publication directly addressed the analysis of the relationship between UDL principles and student achievement. A recent literature review (Rao, Ok, and Bryant, 2014) identified a variety of approaches in terms of how Universal Design (UD) and Universal Instructional Design (UID) were defined and how student outcomes were quantified. The research review identified eight research articles focused on some application of UD/UID in a preK-12 setting.

The seeming lack of research, however, may be a result of the overarching concept of UDL not being addressed, but only pieces or principles of UDL applied to student learning. The National Center on Universal Design for Learning (2012), for example, has cited almost 1,000 research articles that support the different checkpoints and principles of UDL. As an example, under the Provide Multiple Means of Engagement principle, checkpoint 8.3 *Foster collaboration and community* has over 40 citations of experimental and quantitative evidence, or research studies, to support the design element.

**Social Justice and Student Achievement.** The connection between socially just teaching practices and student achievement has now been discussed in terms of a OTL and UDL perspective. Both OTL and UDL overlap with social justice because of the

emphasis all three concepts have on equity, access and fairness. The direct connection between socially just teaching practices and student achievement (Mitescu, Cochran-Smith, Pedulla, Cannady, and Jong, 2011), and the evaluation of the impact of socially just teaching practices (Dover, 2009) are two gaps in the current literature. There is research that explores the use of teaching for social justice and its effects on students. Gutstein (2007), for example conducted a qualitative study identifying and describing ways that teaching for social justice increased student agency in middle school math classrooms. In addition, Mitescu et al. identified a correlation between teaching practices associated with social justice and student achievement, there are a number of limitations to generalizability noted by the authors. While the variables socially just teaching practices and student math achievement accounted for 19% of the variance in the sample of 22 novice teachers, the sample was not random and causation could not be inferred. The sample consisted of first and second year teachers from a variety of grades, ranging from first through sixth. To assess math achievement, an end of unit math assessment was used, but the assessments could not be equated across grade levels (Mitescu et al.)

This present study addresses many of the limitations in Mitescu et al. The main contribution of this study is the ability to randomly assign class lists to teachers. A second point of value added by this study is the ability to have a standardized dependent variables in the regression models used.

This study also adds to the current body of research by examining the relationship between socially just teaching practices for traditionally underserved populations of



students. This study specifically analyzes the relationship between socially just teaching practices and particular groups of marginalized students, because it is hypothesized that socially just teaching practices will have a positive and significant relationship with student achievement for each of the subgroups.

There is research documenting the achievement gap for ELLs (Echevarria, Vogt, and Short, 2013; Fry, 2007), SWDs (Schulte, Stevens, Elliott, Tindall and Nese, 2016) and students in poverty (Nichols et al., 2012). Achievement gaps in reading and mathematics for Black (Vanneman, Hamilton, Anderson, and Rahman, 2009) and Hispanic (Hemphill, Vanneman, and Rahman, 2011) students relative to white students have also been documented. This study will analyze the relationship between socially just teaching practices and achievement for Black, Hispanic or American Indian students. As Ladson-Billings (2006) argues, the view of achievement gap can also be understood as education debt that has built up for these subgroups of students. Socially just teaching is intended to reduce some of the debt described by Ladson-Billings.

### **Chapter 3 - Methodology**

This study is a secondary analysis of the Measures of Effective Teaching Database. A subsample of the 4th and 5th grade English Language Arts (ELA) teachers, and their students, who participated in the Measures of Effective Teaching Project was selected. The Measures of Effective Teaching Project includes data on over 40,000 students and their approximately 3,000 teachers from the following public school districts: Charlotte-Mecklenburg Schools, N.C.; Dallas Independent School District, Texas; Denver Public Schools, Colo.; Hillsborough County Public Schools, Fla.; Memphis City Schools, Tenn.; New York City Department of Education, N.Y (Bill and Melinda Gates Foundation, 2013). A sub-sample of English Language Arts classroom teachers from grades 4 and 5 classroom was taken for this study. Tables 2 and 3 (Inter-University Consortium for Political and Social Science Research, n.d.) include a breakdown of the data based on subject and grade, respectively for the entire MET database.

#### **Human Subjects Protection**

The data from the MET project was accessed through the Measures of Effective Teaching Longitudinal Database housed by the Inter-university Consortium for Political and Social Research (ICPSR) housed at the University of Michigan - Ann Arbor. An application to access the database was required, and part of the application included a \$350 access fee and a data use agreement. The user agreement stipulates protections to research participants to ensure that confidentiality of participants in the study are

protected. For example, coefficients in regression analysis were not allowed to be published for district or school identification variables to ensure confidentiality. Prior to obtaining access to the MET longitudinal database through the ICPSR, approval for the study was obtained from the Institutional Review Board at Boston College.

**Table 2**

*Frequency and percentage of participants in MET Project Videos by Subject*

<b>Subject</b>	<b>Frequency</b>	<b>Percentage (%)</b>
ELA	12,029	48.8
Mathematics	11,809	47.9

**Table 3**

*Frequency and percentage of participants in MET Project Videos: by Grade*

<b>Grade</b>	<b>Frequency</b>	<b>Percentage (%)</b>
4	4633	18.8

### **Data Sources and Collection Procedures**

The study was limited to a sub-sample of the 24,659 videos (Inter-University Consortium for Political and Social Science Research, n.d.) that are part of the MET longitudinal data base. The reason for limiting the study to data of teachers and students whose classrooms were videoed as part of the project is the need to score the videos

based on socially just teaching practices using the Teaching for Social Justice Observation Scale (TSJOS) of the Reformed Teaching Observation Protocol-Plus (RTOP+) (Mitescu, Cochran-Smith, Pedulla, Cannady, and Jong, 2011).

The subsample used for this study was originally limited to 4th grade ELA or ELA and Mathematics teachers, for whom there was teacher experience data available, and for whom classroom videos existed in the MET database. This initial sample yielded 97 teachers. This number, however, was determined to not have enough power for the regression analyses in this study, and some of the teachers who had codes for classroom videos in the database in fact did not have videos in the secure streaming video database. Therefore, the decision was made to expand the sub-sample to 4th and 5th grade ELA classrooms, and also remove the inclusion criteria of teachers in the sub-sample having years of experience data available.

This yielded 354 teachers from whom to take a random sample. From the 354 teachers, 100 were randomly selected using a random number generator. When cross-referencing video availability for these 100 teachers, 51 teachers had videos in the secure streaming database. Because the power analysis indicated the study would require approximately 100 teachers to detect significant effects for socially just teaching practices based on the proposed regression models, an additional 100 teachers were randomly selected from the remaining 254 teachers. This sampling produced a final number of 107 teachers, they are included in this study with their 2,587 students. Videos from Year 2 of the MET study will be used because in Year 2 teachers were randomly

assigned classes of students. The random assignment of classes to teachers limits the probability of teacher-level variables other than socially just teaching practices to explain variance in the student achievement outcome variable.

### **Data Sources/Instruments**

The study utilized a number of measures included in the MET project and add measures of socially just teaching practices.

**Student and Teacher Demographics.** Demographic information on teachers and students collected by the MET was included data in this study (see Tables 4 and 5).

Teacher variables included the race, gender, the number of years teaching as well as years in the district, school, grade level and content taught at the time of assessment. Teacher content knowledge, as measured by Content Knowledge for Teaching assessments developed for the MET in ELA, is another variable that used in this study. When analyzing student achievement at the class level, rather than the student level, the z-score for section ELA average on the 2011 standardized state assessment was used.

**Table 4***Teacher Descriptive Statistics for Sub-Sample*

	<b>Frequency/Mean</b>	<b>Valid Percentage/Standard Deviation</b>
<i>Teacher Characteristics</i>		
Male	10	9.3%
Female	93	86.9%
Missing: Gender	4	3.7%
White	79	73.8%
Black or Hispanic	25	23.4%
Missing: Race	3	2.8%
Years of Teaching Experience (n = 59 out of 107 cases))	10.07	8.994
Teacher Knowledge Assessment z-score ELA (n = 92 out of 107 cases)	.2521	.88003
State Assessment 2011: Section Average Rank based z-score, ELA	.1537	.60017

Student information included grade, and teacher. Student level information was analyzed for grade, gender, English Language Learner status, Special Education status, Gifted/Advanced status, Free or Reduced Price Lunch status, and gifted/advanced status. The variable of race will be coded as 0 = Caucasian or Asian and 1 = African American, Hispanic, or American Indian. The reason for coding the race variable in this manner is African American and Latino students has a history of lower levels of achievement compared to Caucasian students, (Nichols, Glass, and Berliner, 2012; Span and Rivers,

2012). Large scale research (Nichols et al.) looking at the achievement gap compares Caucasian students to African American and Hispanic students, and other research (Alliance for Excellent Education, 2010) analyzes Caucasian and Asian graduation rates to students of color, thus setting a precedent for not including Asian students with African American and Hispanic students.

Student level achievement was measured with two different dependent variables. The rank-based z-score from state ELA exam 2011 was used to measure student achievement using traditional statewide assessments. To measure more complex achievement, the z-score scaled score for the Stanford Achievement Test (SAT9) Open-Ended in ELA for each student was used in this study. The reason for including two different dependent variables was to answer research question number two in this study which looks to examine if the explanatory power of socially just teaching practices is dependent on the type of student achievement measure being employed. In addition, there is evidence (Polikoff, 2014) to suggest that there is variation in state-to-state instructional sensitivity in ELA, and the SAT9 provides a consistent measure of student achievement independent of the state the student is in.

**Table 5***Student Descriptive Statistics for Subsample*

	<b>Frequency/Mean</b>	<b>Valid Percentage/Standard Deviation</b>
<i>Student Characteristics</i>		
Age (years)	9.35	.919
Fourth Grade	1332	51.5%
Fifth Grade	1255	48.5%
Missing: Grade Level	0	0%
Subject: ELA	1045	40.4%
Subject: ELA and Math	1542	59.6%
Missing: Subject	0	0%
Male	1290	48.5%
Female	1255	49.9%
Missing: Gender	42	1.6%
Gifted/Advanced	335	12.9%
Not Gifted/Advanced	2210	85.5%
Missing: Gifted/Advanced	42	1.6%
Special Education	271	10.5%
Not Special Education	2274	87.9%
Missing: Special Education	42	1.6%
English Language Learner	356	13.8%
Not English Language Learner	2189	84.6%



Missing: ELL Status	42	1.6%
Free/reduced-price lunch	936	36.2%
Non Free/reduced price lunch	987	38.2%
Missing: Free/reduced price lunch	664	25.7%
Black, Hispanic or American Indian	1475	57%
Non Black, Hispanic or American Indian	1070	41.4%
Missing: Black, Hispanic or American Indian	42	1.6%
White or Asian	999	38.6%
Non White or Asian	1546	59.8%
Missing: White or Asian	42	1.6%
Rank based z-score from state ELA exam 2010 (n = 2304)	.1792	.96936
Rank based z-score from state ELA exam 2011 (n = 2455)	.1620	1.00334
Student's z-score scaled score on SAT9 (n = 2198)	.1637	.95880

**Student sub-groups.** This study examined the impact of socially just teaching practices on student achievement for all students (research question 1a). In addition, the research examined the impact of socially just teaching practices on different sub-groups of students, namely English Language Learners (research question 1b), students with disabilities (1c), students receiving free or reduced price lunch (1d), and gifted/advanced students (1e), and students from traditionally underserved races (research question 1f).

The rationale for grouping students in this way is the research on the achievement gap for ELLs (Echevarria, Vogt, and Short, 2013; Fry, 2007), SWDs (Schulte, Stevens, Elliott, Tindall and Nese, 2016), and low socioeconomic status (free or reduced price lunch as a proxy) (Nichols et al., 2012) suggests that each of these groups are traditionally underrepresented groups who have systemic lower levels of achievement and opportunities than those students who live in a system of privilege. The gap at which gifted students as achieved compared to all students has been described as stable on one analysis of a statewide reading assessment (Schulte et. al, 2016).

**Student Achievement.** Student achievement was measured in two ways through the MET project; through mandated statewide assessments and additional assessments of higher order thinking skills. The additional assessments are the Stanford 9 Open-Ended Reading Response for grade 4 ELA (Bill and Melinda Gates Foundation, 2010a). There is a total of two outcome variables in this study, both of which measure student achievement: 1. Standardized State ELA Assessments and 2. Stanford 9 Open-Ended Reading Response. There is a total of six different statewide ELA assessments that vary in length and format, as the six districts included in the study are located in different states. The common denominator of the six state assessments is they are given to all public school students, and they measure achievement in English Language Arts. The six state assessments in the MET database had significantly different correlations to teacher VAM scores, and sensitivity to instructional was low overall (Polikoff, 2014). There are two main reasons for including the six statewide ELA assessments in this analysis,

although they differ in their correlation to teacher VAM score. The first reason is practicality, as each state assessment is what is typically used if and when teacher VAMs are used and include student achievement data. In other words, the current political and legal context determines that student achievement is measured and defined by these statewide assessments for terms of school, district and state accountability. Second, although there is variation between the six state assessments, the district and school that each teacher and student is in will be put in any hierarchical regression analyses in this study, which should capture variation due to differences in the state assessment take, as all students in a particular district are in a specific state, and all students in that state take the same statewide ELA assessment. The SAT9 Open Ended ELA assessment includes a narrative text that students read and then answer nine open-ended questions about the text and is intended to measure more cognitively demanding skills like critical thinking (Grossman, Cohen, Ronfeldt, and Brown, 2014).

**Classroom Observation Instruments.** The observations of teaching practices occurred one time during one year for each teacher included in the study. Each classroom observation was 20-22 minutes in length, measured from when the teacher began the lesson. There were two video coders in this study. The author of this study coded videos. The research assistant on this study, who is an experienced elementary teacher and certified reading specialist, also coded videos using the TSJOS in this study. The video coders coded a shared sub-sample of videos, not included in the final analysis, until an inter-rater reliability level of at least 80% on three video observations in a row was

reached. The video coders agreed on 92.8% of items for three videos, with agreement percentages for the three videos of 85.7% (12 out of 14 items in agreement), 93% (13 out of 14 items in agreement) and 100% (14 out of 14 items in agreement), respectively.

Three additional classroom videos were analyzed by both raters after all videos in the study sample had been coded. The coders had an agreement rate of 90.48%, with agreement percentages for the three videos of 93% (13 out of 14 items in agreement) for two of the videos and 86% (12 out of 14 items in agreement) for the third. For both the pre and post inter-rater reliability observations, none of the items in disagreement varied by more than 1 point on the observation scale. A comparison of teacher average TSJOS score and item variability is part of the analysis.

***Socially Just Teaching Practices.*** The subsamples of videos for this proposed study were analyzed using the Teaching for Social Justice Observation Scale (TSJOS) of the Reformed Teaching Observation Protocol-Plus (RTOP+) (Mitescu, Cochran-Smith, Pedulla, Cannady, and Jong, 2011) to measure socially just practices during each lesson. The teacher average score for the TSJOS observation tool is measurement of socially just teaching practices for this study. The TSJOS consists of fourteen items and all items have been found to have internal consistency at a high level (Mitescu et al., 2011). It consists of eight items from the original RTOP as well as six items developed by Mitescu et al. that captures concepts related to socially just teaching practices not included in the original RTOP. The items are scales, with raters giving a score between 0 and 4 for each item. A score of 0 indicates no evidence of the item being observed and a score of 4

indicates the observations is very descriptive of the item. As Mitescu et al. note, a score of 2 is indicative as substantial evidence of a particular socially just teaching practice being observed.

***Framework for Teaching.*** The MET database includes at least four videos of each teacher in the database. It also includes scores from five different classroom observation tools applied to those videos by the MET; one of those five tools is the Framework for Teaching (FtF) (Danielson, 2007). For the purposes of this study, select subscores from the FtF that are aligned to Universal Design for Learning principles were used. The specific subsections of the FtF are identified using a crosswalk document that compares UDL and the FtF (Cast and The Danielson Group, 2014). The hypothesis is classrooms that demonstrate a higher level of average FtF scores consisting of subsections that are aligned with UDL principles will be associated with higher gains in student achievement. In addition, the study analyzed the relationship between socially just teaching practices and student achievement after accounting for average FfT scores. Making the argument that FfT scores can serve as a proxy for UDL practices in a classroom, then if socially just teaching practices have a significant relationship with student achievement after accounting for FfT scores, there is limited evidence that socially just teaching is a unique construct from UDL, or at the very least the constructs of effective teaching measured in the FfT.

It is important to note that the FtF is not a validated measure of UDL, there is only a proposed alignment between the FtF and UDL, as presented by experts in each area

(CAST and the Danielson Group, 2014). The MET coded the classroom videos by several FtF components, which fall under one of four FtF domains: Planning and Preparation, The Classroom Environment, Instruction, and Professional Responsibility. Eight FtF components included in the video coding align with the UDL principles, they correspond to two domains, The Classroom Environment and Instruction (see Table 6 for FtF components included in the MET coding scheme as well as alignment of each component to UDL principles). Each video in the MET longitudinal database includes one score for each of the eight components based on 20-22 minutes of observation for all components.

**Table 6**

*FtF Components included in MET Video Scoring and Alignment to UDL Principles\**

<b>FtF Domain</b>	<b>FtF Component</b>	<b>Alignment to UDL Principle(s) (Provide multiple means of ...)</b>
<i>Domain 2: The Classroom Environment</i>	2a: Creating an Environment of Rapport and Respect	Engagement
	2b: Establishing a Culture for Learning	Action and Expression Engagement
	2c: Managing Classroom Procedures	Action and Expression and Engagement
	2d: Managing Student Behavior	Action and Expression and Engagement
<i>Domain 3: Instruction</i>	3a: Communicating with Students	Representation, Action and Expression and Engagement
	3b: Using Questioning and Discussion Techniques	Representation, Action and Expression and Engagement
	3c: Engaging Students in Learning	Representation, Action and Expression and Engagement
	3d: Using Assessment in Instruction	Representation, Action and Expression and

		Engagement
--	--	------------

\*Table created based on information from Danielson (2007) and CAST and The Danielson Group (2014).

***Protocol for Language Arts Teaching Observation (PLATO).*** Average observation scores from the Protocol for Language Arts Teaching (PLATO) was used in this study to address research question #3 around the explanatory power of socially just teaching observation scores above and beyond other observational measures of teacher effectiveness. The MET project included ratings on seven sub-sections of the PLATO observation scale, and the average PLATO score across all observations that were completed during the MET project for each of the 7 items is in the database as a separate variable for each teacher included in this study was used to predict student achievement scores. The subsections of the PLATO observation tool are: Intellectual Challenge, Classroom Discourse, Behavior Management, Modeling, Strategy Use & Instruction, Time Management, and Representation of Content. In the MET study raters were trained by Educational Testing Services (ETS) on the observation protocols used to rate classroom videos in the MET database (Bill and Melinda Gates Foundation, 2010b).

### **Data Analysis Plan**

The data analysis includes a thorough investigation of demographic information to further describe the subsample used in the study. The sub-sample was randomly selected from the MET database, and was limited to fourth and fifth grade ELA teachers.

The data analysis in this study consisted of thirteen different regression models (see Pedhazur, 1997). All of the regression analyses examined the relationship between

socially just teaching practices, as measured by teacher mean TSJOS scores, and a measure of student achievement (state exam or SAT9). Assumptions for each regression model were assessed for each regression model, and a summary of evidence supporting those assumptions is included in the findings chapter of this paper. The following statistics are reported in the findings section for each regression: coefficients and standardized coefficients for each variable entered in the model (note: all coefficients for the variables District ID and School ID are not reported or released per the MET Data Use Agreement)  $R$ ,  $R$ -squared, Adjusted  $R$ -Square,  $R$ -square change (for hierarchical multiple regression analyses), and  $F$  change. The significance of the slope of the independent variable(s) are also addressed, with a benchmark of  $p < 0.05$  as defining significance for this study for each regression coefficient. The significance of the  $F$  statistic, and  $F$  change statistic for hierarchical regression analyses, will also be reported, with  $p < .05$  defining significance.

Table 7 summarizes the regression models and includes the variables included in each regression model. The column with the heading “Data Split?” indicates whether or not the subsample was parsed into separate groups for that particular regression. Regression 1, for example, included an analysis of the entire subsample, while Regression #4 consisted of a regression analysis for students who were classified as ELLs and a separate analysis consisting of students who were not classified as ELLs. The splits were done in order to run a linear regression that allowed the study to determine the significance of meant teacher TSJOS score to predict student achievement scores



specifically for the subgroup (ELL, SWD, etc.) of interest.

**Table 7**

*Summary of Regression Models*

R#	RQ	Type	Data Split?	DV	Model 1	Model 2	Model 3	Model 4
1.	1a.	Unadjusted Regression Model	No	Z-score section average from state ELA exam 2011	Mean TSJOS score	X	X	X
2.	1a.	Unadjusted Regression Model	No	Student rank based z-score from state ELA exam 2011	Mean TSJOS score	X	X	X
3.	1a.	Hierarchical multiple regression	No	Student rank based z-score from state ELA exam 2011	Location variables *	Teacher Variables*	Student Variables ***	Mean TSJOS Score
4.	1a.	Unadjusted Regression Model	Yes - by grade	Z-score section average from state ELA exam 2011	Mean TSJOS	x	X	X
5.	1a.	Unadjusted Regression Model	Yes - by grade	Student z-score scaled score on SAT9)	Mean TSJOS	x	X	X
6.	1b.	Unadjusted Regression Model	Yes - based on student ELL status	Student rank based z-score from state exam 2011	Mean TSJOS score	x	X	X
7.	1c.	Unadjusted Regression Model	Yes - based on student SWD status	Student rank based z-score from state exam 2011	Mean TSJOS score	x	X	X
8.	1d.	Unadjusted Regression Model	Yes - based on student	Student rank based z-score from state exam 2011	Mean TSJOS score	x	x	x

			nt free/r educed price lunch status					
9.	1e.	Unadjusted Regression Model	Yes - based on stude nt Gifted /Adva nced status	Student rank based z-score from state exam 2011	Mean TSJOS score	x	X	X
10.	1f.	Unadjusted Regression Model	Yes - based on stude nt race status ****	Student rank based z-score from state exam 2011	Mean TSJOS score	x	X	X
11.	2	Unadjusted Regression Model	No	Student z-score scaled score on SAT9	Mean TSJOS score	x	X	X
12.	3	Hierarchic al multiple regression	No	z-score section average from 2011 ELA state exam	Average FtF2 ELA scores	Mean TSJOS score	X	X
13.	3	Hierarchic al multiple regression	No	z-score section average from 2011 ELA state exam	Teacher variables	Average FtF average ELA subscores	Mean TSJOS score	X
14.	3	Hierarchic al multiple regression	No	z-score section average from 2011 ELA state exam	Average PLATO scores	Mean TSJOS score	X	X
15.	3	Hierarchic al multiple regression	No	z-score section average from 2011 ELA state exam	Teacher variables	Average PLATO score	Mean TSJOS score	X

R# = Regression number

RQ = Research Question

DV = Dependent variable

\* Location Variables = School ID; District ID

\*\*Teacher Variables = Gender; Race; Knowledge Assessment Score

\*\*\*Student Variables = Gender; Race (White or Asian); Race (Black, Hispanic or American Indian); ELL Status; SWD Status; Free/Reduced price lunch status;

Gifted/Advanced status; z-score from 2010 ELA state exam

\*\*\*\*Race variable indicating if student is Black, Hispanic or American Indian was used to split the data in this regression analysis.

The following is a summary of the regression analyses described in Table 8 organized by research question (and sub-question).

Research Question 1a. What is the relationship between teacher Social Justice classroom observation scores and student achievement for all students? Regression models 1-5 address this research question. Model 1 is a simple linear regression with mean TSJOS score as the independent variable and z-score for section average on the 2011 ELA state exam as the dependent variable. In this analysis, student achievement is measured at the class, or section, level, rather than by student achievement for each individual student.

Model 2 is the same simple linear regression as Model 1, with the exception that the dependent variable measuring student achievement is each student's rank based z-score from the 2011 ELA state exam.

Model 3 is a hierarchical multiple regression with student rank based z-score on the 2011 ELA state exam as the dependent variable. The independent variables in this regression were forced in by the author. The first block of predictor variables consisted of location variables, namely the district and school of the student. The second block of predictor variables were teacher variables, followed by a group of student variables. The decision was made to exclude the teacher variable measuring years of teaching experience because 45% (48 out of 107) of the teachers had missing data for that variable. The fourth and final block of predictor variables in this model was the teacher mean TSJOS score. The reason for the setup of this model was to determine if the measurement of teachers' socially just teaching practices has a significant relationship with student achievement after accounting for location, teacher and student predictor variables, including the student's prior year score on the state ELA exam. All of these predictor variables are hypothesized to explain some proportion of variance in student achievement scores.

Models 4 and 5 splits the data by grade level and then ran a simple linear regression that predicted student achievement from teacher mean TSJOS score. Model 4 uses the section rank based z-score on the 2011 ELA state exam as the dependent variable, and Model 6 uses the student z-score scaled score from the SAT9 as the dependent variable.

Research Question 1b., c., d., e., and f. What is the relationship between teacher Social Justice classroom observation scores and student achievement for ...

- b. English Language Learners (and non-ELL)?
- c. Students with Disabilities (SWD) (and non-SWD)?
- d. Students receiving free or reduced price lunch? (and students who do not receive free/reduced price lunch)?
- e. Gifted/Advanced students (and non-gifted/advanced students)?
- f. Students who are Black, Hispanic, or American Indian (and students who are not Black, Hispanic or American Indian)?

Regression analyses 6-10 as outlined in Table 8 are all simple linear regressions With a dependent variable of student rank based z-score from the 2011 ELA state exam. The independent variable for each of these analyses is the mean TSJOS score for the teacher of each student. For each regression, the data was split based on the sub-group of interest. For example, in model 6, the data was split based on the variable indicating the ELL status of each student. These analyses yielded statistics for each sub-group as well as for all students who were not part of that sub-group. For example, the statistics in model 6 allow for an analysis of the relationship between socially just teaching practices and student achievement for both ELL students and Non-ELL students.

Research Question 2: What is the impact of the type of student achievement measure (standardized state tests verses tests that measure higher-order thinking) on the relationship between teacher social justice classroom observation scores and student achievement? Regression #11 in Table 8 addresses this research question, where the dependent variable is the student's z-score scaled score on the SAT9 and the independent

variable is the teacher mean TSJOS score. A comparison between the correlation coefficient for this model is made with the simple linear regression with the dependent variable of student rank based z-score on the 2011 ELA state exam (regression #2 in Table 8).

Research Question 3: What, if any additional explanation of variance in student achievement scores can be explained by teacher social justice classroom observation scores above and beyond the variance in scores explained by (i.) Average Framework for Teaching (FfT2) ELA observation scores and (ii.) Average Protocol for Language Arts Teaching Observation (PLATO) scores. Regressions 12-15 in Table 8 address this research question through hierarchical multiple regression, and all four analyses have the z-score section average on the 2011 ELA state exam as the dependent variable measuring student achievement. Regression #12 and #13 focus on teacher FfT observation scores. Regression #12 includes the average ELA score for each of the eight FfT2 sub-scales for the first model and then has the teacher mean TSJOS as the second and final model. Regression #13 in Table 8 enters a set of teacher variables (race, gender and knowledge assessment score) as the first model, and follows with the same FfT2 scores as analysis #12 for the second model and mean TSJOS as the third model. The intent of these analysis is to determine the relationship between socially just teaching practices and student achievement after accounting for teacher variables and FfT observational scores.

Regression analyses #14 and #15 in Table 8 focus on teacher PLATO observation scores. Analysis #15 entered in the seven PLATO subscale variables as the first model

and follows with teacher mean TSJOS as the second model. Analysis #15, in a similar setup to analysis #13, entered in the set of teacher variables into the model first, then added in the PLATO variables second, and ended with teacher mean TSJOS as the third block.

In summary, this study uses a combination of simple linear regression and hierarchical multiple linear regression analyses (see Pedhazur, 1997 and Shavelson, 1996) to examine the relationship between socially just teaching practices and student achievement. The findings from the methods and data analysis described in this chapter follow.

## Chapter 4 - Findings

### Socially Just Teaching Practices

The overarching objective of this study is to describe the relationship between socially just teaching practices and student achievement. For the sample of 107 teachers included in this study, the mean TSJOS score was 2.2991 with a standard deviation of .77381. Mean scores on TSJOS items ranged from 1.97 (item 27. “The instruction strategies accommodate and scaffold the learning of all students, including, when appropriate, boys and girls, ELLs, SPED, and students with diverse racial and cultural backgrounds”) to 2.63 (item 31. “There is a fair and just environment that is promoted by the teacher and embraced by the students.”). Table 8 summarizes the data collected through analysis of classroom videos using the TSJOS observational tool to measure socially just teaching practices.

**Table 8**

*Item Mean Scores and Overall Average on Teaching for Social Justice Observation Scale (TSJOS)*

	N	Mean	Std. Deviation	Variance
TSJOS01	107	2.43	.870	.757
TSJOS02	107	2.31	.926	.857
TSJOS10	107	2.26	.828	.686
TSJOS13	107	2.30	.983	.966
TSJOS15	107	2.14	.905	.820
TSJOS17	107	2.25	.891	.794



TSJOS20	107	2.61	.898	.807
TSJOS22	107	2.21	.855	.731
TSJOS26	107	2.12	.832	.693
TSJOS27	107	1.97	.852	.726
TSJOS28	107	2.11	.915	.836
TSJOS29	107	2.44	.923	.852
TSJOS30	107	2.41	.971	.943
TSJOS31	107	2.63	.864	.746
Mean TSJOS Score	107	2.2991	.77381	.599
Valid N (listwise)	107			

**Findings for Research Question 1a.: What is the relationship between teacher Social Justice classroom observation scores and student achievement for all students?**

The regression analyses conducted provide evidence of a significant relationship between socially just teaching practices and student achievement for all students.

Table 9 summarizes the simple linear regression that was run with the dependent variable of z-score sections average from state 2011 ELA exam as a measure of student achievement. The independent variable was teacher mean TSJOS score. Assumptions of simple linear regression were all met. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 2.16,

suggesting the assumption of independence of residuals was met. The normality assumption was met by visual inspection of the Q-Q plot. There were no significant outliers in the sample, as evidenced by the maximum leverage point being less than 0.2, no Cook's distance values greater than 1, and no cases whose residuals were greater than plus or minus 3 standard deviations.

**Table 9**

*Regression #1: Predicting z-score on section average 2011 ELA state exam from mean TSJOS Score.*

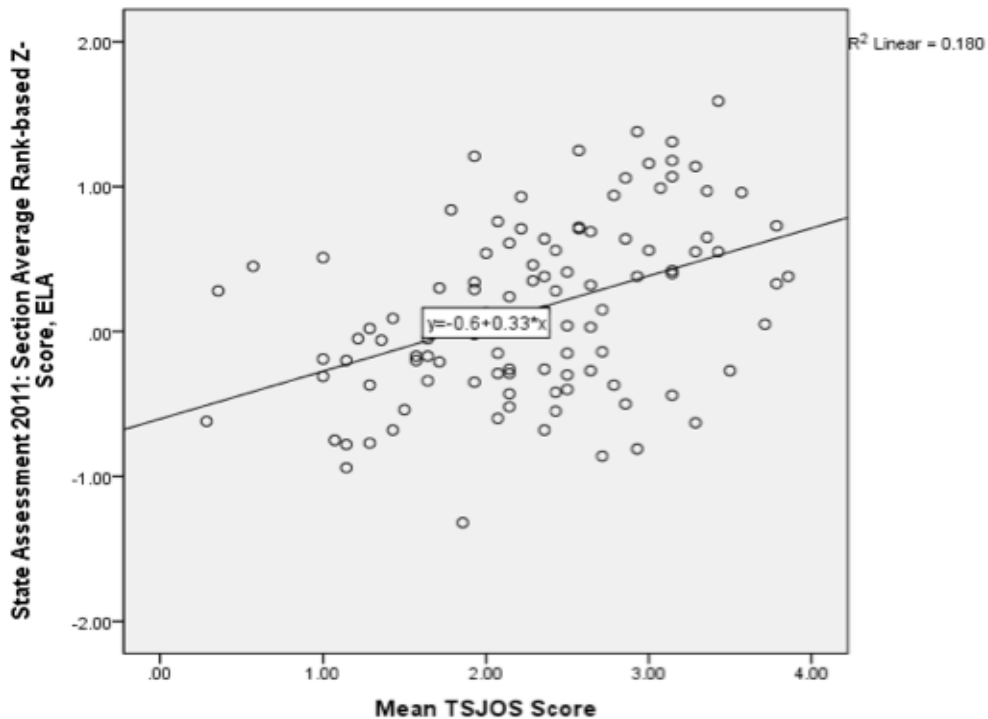
Variable	Model 1	
	B	B
Constant	-.603***	
Mean TSJOS Score	.329***	.425
R	.425	
R-Square	.180	
Adjusted R-Square	.172	
F	23.092***	
<i>Note: N = 107 teachers, *p &lt; .05, **p &lt; .01, ***p &lt; .001</i> Socially Just Teaching = Average Score on TSJOS observation Scale		

Teacher mean TSJOS scores significantly predicted the section average 2011 state ELA exam score, with  $\beta = .425, p < .001$ . Mean TSJOS score also explained a significant

proportion of variance in section 2011 ELA state exam score, with Adjusted  $R^2 = .172$ ,  $F(1, 105) = 23.092$ ,  $p < .001$ . See Figure 2 for a visual representation of this regression analysis.

**Figure 2**

*Student Achievement predicted by Socially Just Teaching Practices*



Regression #2 differs from the first regression analysis in that the dependent variable is predicting individual student exam scores rather than the section average for each teacher. Table 10 summarizes the simple linear regression that was run with the dependent variable of student rank based z-score from state 2011 ELA exam as a measure of student achievement. The dependent variable was teacher mean TSJOS score.

Assumptions of simple linear regression were all met, with the exception of a Durbin-Watson value that could indicate possible auto-correlation. Results of this regression, therefore, should be interpreted with caution. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 1.206, suggesting possible auto-correlation and violation of the assumption of independence of residuals. The tolerance value was equal to 1, suggesting there was no multicollinearity. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point being less than 0.2 and no Cook’s distance values greater than 1, suggesting there were no influential outliers in this analysis. There were six cases whose residuals were greater than plus or minus 3 standard deviations. The removal of those six cases had no effect on the significance of the regression analysis and the decision was made to keep the cases in the analysis.

**Table 10**

*Regression #2: Predicting student rank-based z-score on 2011 state ELA exam from Mean TSJOS Score*

Variable	Model 1	
	B	B
Constant	-.632***	
Mean TSJOS Score	.344***	.265

R	.265	
R-Square	.07	
Adjusted R-Square	.07	
F	185.634***	
<i>Note:</i> $N = 2455$ students, * $p < .05$ , ** $p < .01$ , *** $p < .001$ Socially Just Teaching = Average Score on TSJOS observation Scale		

Teacher mean TSJOS scores significantly predicted student rank based z-score from the 2011 ELA state exam, with  $\beta = .265$ ,  $p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .07$ ,  $F(1, 2453) = 185.634$ ,  $p < .001$ .

After establishing a significant relationship between socially just teaching and student achievement, a hierarchical multiple regression was conducted that entered in location, teacher and student variables as the first three models, respectively, before adding in teacher Mean TSJOS score (Model 4) (see Table 11). Assumptions of hierarchical multiple regression were all met. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic for the final model (Model 4) was equal to 1.918, suggesting the assumption of independence of residuals was met. None of the independent variables in the mode had a tolerance value less than 0.1, suggesting there was no multicollinearity. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point being less than 0.2 and no

Cook's distance values greater than 1, suggesting there were no influential outliers in this analysis. There were 11 cases whose residuals were greater than plus or minus 3 standard deviations. The removal of those 11 cases had no effect on the significance of the regression analysis and the decision was made to keep the cases in the analysis.

**Table 11**

*Regression #3: Predicting student achievement for all students from location, teacher, student and socially Just Teaching Variables*

Variable	Model 1		Model 2		Model 3		Model 4	
	B	$\beta$	B	$\beta$	B	$\beta$	B	B
Constant	.679***		.506*		.039		-.176	
District ID	x	X	x	X	X	X	x	X
School ID	x	X	X	x	X	x	X	X
Teacher Gender			-.007***	-.112	-.001	-.022	-.001	-.018
Teacher Race			.000	-.007	-.001	-.016	-.001	-.012
Teacher Knowledge Assessment Score			.197***	.189	.026	.025	.023	.022
Student Gender					.001	.001	.000	.000
Student Gifted					.346***	.128	.335***	.124
Student Special Education Status					-.166**	-.051	-.164**	-.051
Student ELL					.031	.009	.026	.008

Status								
Student free-reduced price lunch Status					-.085*	-.042	-.078*	-.039
Student Race – White or Asian					.028	.014	.026	.013
Student Race – Black, Hispanic or American Indian					-.073	-.037	-.074	-.037
Student z-score from state ELA exam 2011					.703***	.688	.700***	.685
Mean TSJOS Score							.050*	.020
R	.240		.323		.821		.821	
R-Square	.057		.105		.673		.675	
Adjusted R-Square	.056		.101		.670		.671	
R-Square Change	.057		.047		.569		.001	
F Change	43.906***		25.273***		311.288***		5.979*	
Note: $N = 1445$ students * $p < .05$ , ** $p < .01$ , *** $p < .001$ X = Coefficients for School and District not released per MET Data Use Agreement								

Adding mean TSJOS to location, teacher and student variables resulted in a statistically significant increase in the proportion of variance in the dependent variable being explained, with a  $R^2$  of .001,  $F = 5.979$ ,  $p = .015$ . Mean TSJOS significantly

predicted student achievement,  $\beta = .020$ ,  $p = .015$ .

Regression #4 first split the data by grade level prior to running a simple linear regression predicting section average z-score on the 2011 ELA state exam from teacher mean TSJOS (see Table 12). Assumptions of simple linear regression were all met for both 4th and 5th grade classrooms. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was 2.33 for the 4th grade subsample and 1.716 for the 5th grade subsample, suggesting independence of residuals. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point being less than 0.2 and no Cook's distance values greater than 1, suggesting there were no influential outliers in this analysis for both 4th and 5th grade analyses.

**Table 12**

*Regression #4: Predicting student achievement for 4<sup>th</sup> grade and 5<sup>th</sup> grade classrooms from socially just teaching scores*

Variable	4 <sup>th</sup> Grade		5 <sup>th</sup> Grade	
	B	$\beta$	B	B
Constant	-.619*		-.631**	
Mean TSJOS Score	.313**	.388	.370***	.490
R	.388		.490	
R-Square	.151		.240	
Adjusted R-Square	.135		.225	



R-Square Change	.151		.240	
F Change	9.776**		15.196***	
	N =57 teachers		N = 50 teachers	

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

Teacher mean TSJOS scores significantly predicted section average z-score from the 2011 ELA state exam, for 4th grade classrooms, with  $\beta = .313$ ,  $p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .135$ ,  $F = 9.776$ ,  $p < .01$ . Teacher mean TSJOS scores also significantly predicted section average z-score from the 2011 ELA state exam for 5th grade classrooms, with  $\beta = .370$ ,  $p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .240$ ,  $F = 15.196$ ,  $p < .001$ .

Regression #6 also split the data by grade level prior to running a simple linear regression, but in this case the analysis predicted student z-score scaled score of the SAT9 from teacher mean TSJOS (see Table 13). Assumptions of simple linear regression were met for both 4th and 5th grade classrooms, with the exception of independence of residuals for both subsamples in this analysis. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was 1.104 for the 4th grade subsample and 1.252 for the 5th grade subsample, suggesting a violation of the assumption of independence of residuals. Therefore, results of these analyses should be interpreted with caution. The normality assumption was met by visual inspection of the

Q-Q plot. The maximum leverage point being less than 0.2 and no Cook's distance values greater than 1, suggesting there were no influential outliers in this analysis for both 4th and 5th grade analyses.

**Table 13**

*Regression #5: Predicting student achievement scores (SAT9) for 4th and 5th grade classrooms from socially just teaching scores*

Variable	4 <sup>th</sup> Grade		5 <sup>th</sup> Grade	
	B	$\beta$	B	B
Constant	-.483***		-.453***	
Mean TSJOS Score	.260***	.21 2	.293***	.225
R	.212		.225	
R-Square	.045		.051	
Adjusted R-Square	.044		.050	
R-Square Change	.045		.051	
F Change	52.112***		57.865***	
	N = 1111 students		N = 1087 students	

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

Teacher mean TSJOS scores significantly predicted z-score scaled score for the SAT9, for 4th grade classrooms, with  $\beta = .212$ ,  $p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .044$ ,  $F = 52.112$ ,  $p < .001$ . Teacher mean TSJOS scores also significantly predicted z-

score scaled score for the SAT9 for 5th grade classrooms, with  $\beta = .225, p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .050, F = 57.865, p < .001$ .

**Findings for Research Question 1b: What is the relationship between teacher Social Justice classroom observation scores and student achievement for English Language Learners (and non-ELL)?**

A simple linear regression was run after splitting the data file based on the student variable indicating if the student had ELL status to determine the prediction power of mean TSJOS on student rank based z score on the 2011 ELA state exam (see Table 14).

Assumptions of simple linear regression were met for ELL students, with the exception of a Durbin-Watson statistic that could indicate dependence of residuals. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 1.606, suggesting the assumption of independence of residuals may not have been met. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point being less than 0.2 and no Cook's distance values greater than 1, suggesting there were no influential outliers in this analysis. There were 3 cases whose residuals were greater than plus or minus 3 standard deviations. The removal of those 3 cases had no effect on the significance of the regression analysis and the decision was made to keep the cases in the analysis.

**Table 14**

*Regression #6: Predicting student achievement for ELLs from socially just teaching scores.*

Variable	ELL		Non-ELL	
	B	$\beta$	B	$\beta$
Constant	-.795		-.596***	
Mean TSJOS Score	.158*	.131	.371***	.297
R	.131		.297	
R-Square	.017		.088	
Adjusted R-Square	.014		.088	
R-Square Change	.017		.088	
F Change	5.889*		203.381***	
	N = 341 ELL Students		N = 2108 Non-ELL Students	

Teacher mean TSJOS scores significantly predicted student rank based z-score for the 2011 state ELA exam score for ELL students, with  $\beta = .131$ ,  $p < .05$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .014$ ,  $F(1, 339) = 5.889$ ,  $p < .05$ .

The same assumptions of simple linear regression were also analyzed for the subsample of Non-ELL students. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 1.434, suggesting the assumption of independence of residuals may not have been met. The results for Non-

ELL students, therefore, should be interpreted with caution. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point being less than 0.2 and no Cook's distance values greater than 1, suggesting there were no influential outliers in this analysis. There were 3 cases whose residuals were greater than plus or minus 3 standard deviations. The removal of those 3 cases had no effect on the significance of the regression analysis and the decision was made to keep the cases in the analysis.

Teacher mean TSJOS scores significantly predicted student rank based z-score for the 2011 state ELA exam score for Non-ELL students, with  $\beta = .297, p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .088, F(1, 2106) = 203.381, p < .001$ .

**Findings for Research Question 1c: What is the relationship between teacher Social Justice classroom observation scores and student achievement for Students with Disabilities (and Non-SWD).**

A simple linear regression was run after splitting the data file based on the student variable indicating if the student had Special Education status to determine the prediction power of mean TSJOS on student rank based z score on the 2011 ELA state exam (see Table 15).

Assumptions of simple linear regression were all met for SWD students. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson

statistic was equal to 1.957, suggesting the assumption of independence of residuals was met. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point was 0.26 and no Cook's distance values greater than 1, suggesting there were no influential outliers in this analysis. There were no cases whose residuals were greater than plus or minus 3 standard deviations.

**Table 15**

*Regression #7: Predicting student achievement for SWDs from socially just teaching scores.*

Variable	Students with Disabilities		Non-Students with Disabilities	
	B	$\beta$	B	B
Constant	-.807		-.623	
Mean TSJOS Score	.124	.086	.369	.291***
R	.097		.291	
R-Square	.010		.085	
Adjusted R-Square	.005		.084	
R-Square Change	.010		.085	
F Change	2.082		207.259***	
	N = 219 SWD		N = 2236 Non-SWD	

Mean TSJOS was not a statistically significant predictor of student achievement for students with disabilities in this study.

Assumptions of simple linear regression for Non-SWD students was conducted. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 1.389, suggesting possible violation of the assumption of residual independence. Therefore, results should be interpreted with caution. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point was less than .2 and no Cook's distance values greater than 1, suggesting there were no influential outliers in this analysis. There were 2 cases whose residuals were greater than plus or minus 3 standard deviations. The decision was made to keep these 2 cases in the analysis as removing the cases had no effect on the significance of the findings.

Teacher mean TSJOS scores significantly predicted student rank based z-score for the 2011 state ELA exam score for Non-SWD students, with  $\beta = .291, p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .084, F(1, 2229) = 207.259, p < .001$ .

**Findings for Research Question 1d: What is the relationship between teacher Social Justice classroom observation scores and student achievement for students that receive free/reduced priced lunch (and students who do not qualify)?**

A simple linear regression was run after splitting the data file based on the student variable indicating if the student had free/reduced price lunch status to determine the prediction power of mean TSJOS on student rank based z score on the 2011 ELA state exam (see Table 16).

Assumptions of simple linear regression were met, with the exception of independence of residuals for students with free/reduced price lunch and for students without free or reduced price lunch. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 1.586, and 1.522 for students who did not qualify for free/reduced price lunch, suggesting the assumption of independence of residuals may have not been met for either group of students. Therefore, the results should be interpreted with caution. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point being less than 0.2 and no Cook's distance values greater than 1, suggesting there were no influential outliers in this analysis. There were 5 cases whose residuals were greater than plus or minus 3 standard deviations, and 1 such case for students who did not qualify for free/reduced price lunch. The removal of those 5 cases had no effect on the significance of the regression analysis and the decision was made to keep the cases in the analysis. For students without free/reduced price lunch, the removal of the 1 case also had no significant effect.

**Table 16**



*Regression #8: Predicting student achievement for students that qualify for free/reduced priced lunch from socially just teaching scores*

Variable	Free/Reduced Price Lunch		Non-Free/Reduced Price Lunch	
	B	$\beta$	B	$\beta$
Constant	-.513***		-.161***	
Mean TSJOS Score	.195***	.041	.274***	.228
R	.156		.228	
R-Square	.024		.052	
Adjusted R-Square	.023		.051	
R-Square Change	.024		.052	
F Change	22.916***		52.088***	
	N = 916 students		N = 949 students	

Teacher mean TSJOS scores significantly predicted student rank based z-score for the 2011 state ELA exam score for students with free/reduced price lunch status, with  $\beta = .041$ ,  $p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .024$ ,  $F(1, 914) = 22.916$ ,  $p < .001$ .

Teacher mean TSJOS scores significantly predicted student rank based z-score for the 2011 state ELA exam score for students without free/reduced price lunch status, with

$\beta = .228, p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .051, F(1, 947) = 52.088, p < .001$ .

**Findings for Research Question 1e: What is the relationship between socially just teaching practices and students achievement for Gifted/Advanced students?**

A simple linear regression was run after splitting the data file based on the variable indicating if a student had Gifted/Advanced status (see Table 17)

Assumptions of simple linear regression were met, with the exception of independence of residuals for students with Gifted status and for those without. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 1.1.642, and 1.540 for students who did not have Gifted status suggesting the assumption of independence of residuals may have not been met for either group of students. Therefore, the results should be interpreted with caution. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point for students with gifted status was .034, suggesting a possible influential point, and no cases in the Non-Gifted group that had a leverage value greater than .2. There were no Cook's distance values greater than 1 for either group of students. There were 7 cases whose residuals were greater than plus or minus 3 standard deviations in the Non-Gifted group of students. The removal of those 7 cases had no effect on the significance of the regression analysis and the decision was made to keep the cases in the analysis. There were no such cases in group for students with Gifted status.

**Table 17**

*Regression #9: Predicting student achievement for gifted/advanced students from socially just teaching scores.*

Variable	Gifted/Advanced		Non- Gifted/Advanced	
	B	$\beta$	B	B
Constant	.956***		-.553***	
Mean TSJOS Score	.140*	.143	.237***	.198
R	.143		.198	
R-Square	.020		.039	
Adjusted R-Square	.017		.039	
R-Square Change	.020		.039	
F Change	6.801*		86.318***	
	<i>N</i> = 330 students		<i>N</i> = 2125 students	

Teacher mean TSJOS scores significantly predicted student rank based z-score for the 2011 state ELA exam score for students with Gifted status, with  $\beta = .143$ ,  $p < .001$ .

Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .02$ ,  $F(1, 328) = 6.801$ ,  $p < .05$ .

Teacher mean TSJOS scores significantly predicted student rank based z-score for the 2011 state ELA exam score for students without Gifted status, with  $\beta = .198$ ,  $p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the

dependent variable, with Adjusted  $R^2 = .039$ ,  $F(1, 2123) = 86.318$ ,  $p < .001$ .

**Findings for Research Question 1f: What is the relationship between socially just teaching practices and student achievement for students who are Black, Hispanic, or American Indian (and students who are not Black, Hispanic or American Indian)?**

A simple linear regression was run after splitting the data file based on the variable indicating if a student was Black, Hispanic or American Indian (see Table 18)

Assumptions of simple linear regression were met, with the exception of independence of residuals for students with minority status and for those without. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 1.436, and 1.292 for students who were not Black, Hispanic or American Indian, suggesting the assumption of independence of residuals may have not been met for either group of students. Therefore, the results should be interpreted with caution. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point for each group of students was less than .2. There were no Cook's distance values greater than 1 for either group of students. There were 4 cases whose residuals were greater than plus or minus 3 standard deviations in the Black, Hispanic or American Indian group of students, and 5 such cases for students who were not Black, Hispanic or American Indian. The removal of those cases had no effect on the significance of the regression analysis and the decision was made to keep the cases in the analysis.

**Table 18**

*Regression #10: Predicting student achievement for Black, Hispanic or American Indian students from socially just teaching scores*

Variable	Minority		Non-Minority	
	B	$\beta$	B	B
Constant	-.673***		-.129	
Mean TSJOS Score	.234***	.191	.296***	.242
R	.191		.242	
R-Square	.037		.059	
Adjusted R-Square	.036		.058	
R-Square Change	.037		.059	
F Change	53.897***		64.437***	
	N = 1419 students		N = 1036 students	

Teacher mean TSJOS scores significantly predicted student rank based z-score for the 2011 state ELA exam score for students who are Black, Hispanic or American Indian with  $\beta = .191$ ,  $p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .036$ ,  $F(1, 1417) = 53.897$ ,  $p < .001$ .

Teacher mean TSJOS scores significantly predicted student rank based z-score for the 2011 state ELA exam score for students who are not Black, Hispanic or American Indian, with  $\beta = .242$ ,  $p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .058$ ,  $F(1, 1034) =$

64.437,  $p < .001$ .

**Findings for Research Question 2: What is the impact of the type of student achievement measure (standardized state tests verses tests that measure higher-order thinking) on the relationship between teacher social justice classroom observation scores and student achievement?**

A simple linear regression was conducted to examine the relationship of teacher TSJOS scores with the dependent variable of student z-scored scaled score on the SAT9 as a measurement of student achievement. The SAT9 is proposed to assess higher-level cognitive processes than assessed on statewide exams (see Table 19).

Assumptions of simple linear regression were all met. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 1.827, suggesting the assumption of independence of residuals was met. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point was less than 0.2, and no Cook's distance values were greater than 1. There were 6 cases whose residuals were greater than plus or minus 3 standard deviations. The decision was made to keep these 6 cases in the analysis because the removal of the cases did not affect the significance of findings.

**Table 19**

*Regression #11: Predicting student z-score scaled scores on SAT9 from mean TSJOS*

scores

Variable	Model 1	
	B	B
Constant	-.451***	
Mean TSJOS Score	.268***	.214
R	.214	
R-Square	.046	
Adjusted R-Square	.045	
F	104.922***	
<i>Note:</i> $N = 2198$ students, * $p < .05$ , ** $p < .01$ , *** $p < .001$ Socially Just Teaching = Average Score on TSJOS observation Scale		

Teacher mean TSJOS scores significantly predicted student z-score scaled score on the SAT9, with  $\beta = .214$ ,  $p < .001$ . Mean TSJOS score also explained a significant proportion of variance in the dependent variable, with Adjusted  $R^2 = .045$ ,  $F(1, 2196) = 104.922$ ,  $p < .001$ .

**Findings for Research Question 3: What, if any additional explanation of variance in student achievement scores can be explained by teacher social justice classroom observation scores above and beyond the variance in scores explained by (i.)**

**Average Framework for Teaching (FFT2) ELA observation scores and (ii.) Average Protocol for Language Arts Teaching Observation (PLATO) scores?**

**Framework for Teaching Findings.** A hierarchical multiple regression was run with the average ELA subscale scores for the FFT2 entered in first (Model 1) and followed by teacher mean TSJOS (Model 2). The dependent variable is z-score section average from 2011 state ELA exam (see Table 20).

Assumptions of simple linear regression were met. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 2.010, suggesting the assumption of independence of residuals was met. Tolerance values for all coefficients in the models were all less than 1, and VIF values were less than 10. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point was .387, and no Cook's distance values were greater than 1. There were no cases whose residuals were greater than plus or minus 3 standard deviations.

**Table 20**

*Regression #12: Predicting student achievement from average ELA FfT scores and mean TSJOS scores.*

Variable	Model 1		Model 2	
	B	B	B	$\beta$
Constant	-3.008***		-2.924***	



FfT: Creating an Environment of Respect & Rapport	-2.18	-.109	-.270	-.135
FfT: Communicating with Students	.466	.252	.375	.203
FfT: Establishing a Culture for Learning	.201	.110	.207	.114
FfT: Engaging Students in Learning	-.031	-.020	-.065	-.041
FfT: Managing Classroom Procedures	.425	.183	.448	.193
FfT: Managing Student Behavior	.062	.031	.082	.041
FfT: Using Assessment in Instruction	-.256	-.138	-.219	-.118
FfT: Using Questioning & Discussion Techniques	.529	.260	.376	.185
Mean TSJOS			.235**	.303
R	.491		.565	
R-Square	.241		.319	
Adjusted R-Square	.179		.256	
R-Square Change	.241		.078	
F Change	3.897***		11.044**	
Note: $N = 107$ Teachers * $p < .05$ , ** $p < .01$ , *** $p < .001$				

Adding mean TSJOS to FtF2 average ELA score variables (Model 2) resulted in a statistically significant increase in the proportion of variance in the dependent variable

being explained, with a  $R^2$  change of .078,  $F$  change = 11.044,  $p < .01$ . Mean TSJOS significantly predicted student achievement,  $\beta = .303$ ,  $p < .01$ .

A second regression analysis that included location and teacher variables prior to putting in the FtF2 scores was then conducted (see Table 21).

Assumptions of simple linear regression were met. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 1.976, suggesting the assumption of independence of residuals was met. Tolerance values for all coefficients in the models were all less than 1, and VIF values were less than 10. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point was .471, and no Cook's distance values were greater than 1. There were no cases whose residuals were greater than plus or minus 3 standard deviations.

**Table 21**

*Regression #13: Predicting student achievement from location and teacher variables, average ELA FfT scores, and mean TSJOS scores*

Variable	Model 1		Model 2		Model 3		Model 4	
	B	$\beta$	B	$\beta$	B	B	B	$\beta$
Constant	.946***		.727*		-2.076*		-1.959*	
District ID	x	x	x	x	x	X	x	x
School ID	x	x	x	x	x	X	x	x
Teacher			-.114	-.058	.085	.043	.070	.035

Gender								
Teacher Race			.000	-.003	.000	-.015	.000	.011
TKA			.205**	.304	.196**	.290	.177**	.261
FfT: Creating an Environment of Respect & Rapport					-.387	-.205	-.406	-.215
FfT: Communicating with Students					.330	.188	.297	.169
FfT: Establishing a Culture for Learning					.262	.146	.256	.142
FfT: Engaging Students in Learning					-.157	-.101	-.240	-.155
FfT: Managing Classroom Procedures					.387	.171	.346	.153
FfT: Managing Student Behavior					.147	.077	.162	.085
FfT: Using Assessment in Instruction					.064	.034	.141	.074
FfT: Using Questioning & Discussion Techniques					.301	.155	.179	.092
Mean							.184*	.236

TSJOS Score								
R	.434		.530		.643		.676	
R-Square	.189		.281		.414		.457	
Adjusted R-Square	.170		.238		.312		.354	
R-Square Change	.189		.092		.133		.043	
F Change	10.004***		3.552*		2.119*		5.881*	
Note: $N = 89$ Teachers * $p < .05$ , ** $p < .01$ , *** $p < .001$ X = Coefficients for District ID and School ID variables not released per MET data use agreement.								

Adding mean TSJOS (Model 4) to Location (model 1), teacher variables (Model 2) and FtF2 average ELA score variables (Model 3) resulted in a statistically significant increase in the proportion of variance in the dependent variable being explained, with a  $R^2$  change of .043,  $F$  change = 5.881,  $p < .05$ . Mean TSJOS significantly predicted student achievement,  $\beta = .236$ ,  $p < .05$ .

**Protocol for Language Arts Teaching Observation Findings.** Similar hierarchical multiple regressions using the FtF (Regressions # 10 and 11) were then conducted using the average PLATO subscale scores in place of FtF scores. The average PLATO subscale score variables were entered in the analysis (Model 1) and then teacher mean TSJOS score was entered (Model 2) (see Table 22).

Assumptions of simple linear regression were met. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value,

indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 2.034, suggesting the assumption of independence of residuals was met. Tolerance values for all coefficients in the models were all less than 1, and VIF values were less than 10. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point was .561, and no Cook's distance values were greater than 1. There were no cases whose residuals were greater than plus or minus 3 standard deviations.

**Table 22**

*Regression #14: Predicting student achievement from average PLATO scores and mean TSJOS scores*

Variable	Model 1		Model 2	
	B	$\beta$	B	$\beta$
Constant	-2.259**		-1.286	
PLATO2: Intellectual Challenge	.117	.070	.071	.042
PLATO2: Classroom Discourse	.111	.075	-.052	-.035
PLATO2: Behavior Management	.358	.193	.158	.085
PLATO2: Modeling	-.206	-.163	-.292	-.231
PLATO2: Strategy Use & Instruction	-.066	-.051	.042	.033
PLATO2: Time Management	.332	.153	.302	.139
PLATO2: Representation of Content	-.341	-.027	-.697	-.055

Mean TSJOS			.316***	.408
R	.365		.504	
R-Square	.133		.254	
Adjusted R-Square	.072		.193	
R-Square Change	.133		.120	
F Change	2.175*		15.795***	
Note: $N = 107$ Teachers * $p < .05$ , ** $p < .01$ , *** $p < .001$				

Adding mean TSJOS to PLATO sub score average variables (Model 2) resulted in a statistically significant increase in the proportion of variance in the dependent variable being explained, with a  $R^2$  change of .120,  $F$  change = 15.795,  $p < .001$ . Mean TSJOS significantly predicted student achievement,  $\beta = .408$ ,  $p < .001$ .

A hierarchical multiple regression analysis was then run by first forcing location, teacher and PLATO variables as three blocks of variables prior to entering teacher mean TSJOS score (Table 23).

Assumptions of simple linear regression were met. A visual inspection suggested little to no relationship between the studentized residuals versus the predicted value, indicating the assumption of linearity was met. The Durbin-Watson statistic was equal to 2.145, suggesting the assumption of independence of residuals was met. Tolerance values for all coefficients in the models were all less than 1, and VIF values were less than 10. The normality assumption was met by visual inspection of the Q-Q plot. The maximum leverage point was .6, and no Cook's distance values were greater than 1. There were no

cases whose residuals were greater than plus or minus 3 standard deviations.

**Table 23**

*Regression #15: Predicting student achievement from location and teacher variables, average PLATO scores, and mean TSJOS scores*

Variable	Model 1		Model 2		Model 3		Model 4	
	B	$\beta$	B	$\beta$	B	$\beta$	B	$\beta$
Constant	.946*		.727**		-.615		.082	
District ID	X	x	x	x	x	x	x	x
School ID	X	x	x	x	x	x	x	x
Teacher Gender			-.114	-.058	-.030	-.015	-.049	-.025
Teacher Race			.000	-.003	-.001	-.021	.000	.000
TKA			.205**	.304	.206**	.306	.183**	.271
PLATO: Intellectual Challenge					.433	.257	.305	.181
PLATO: Classroom Discourse					-.176	-.121	-.235	-.161
PLATO: Behavior Management					.132	.076	-.008	-.005
PLATO: Modeling					-.327*	-.263	-.429*	-.345
PLATO: Strategy Use & Instruction					-.151	-.122	-.035	-.028
PLATO: Time Management					.319	.156	.266	.130

PLATO: Representation of Content					.084	.007	-.085	-.007
Mean TSJOS Score							.261**	.336
R	.434		.530		.652		.705	
R-Square	.189		.281		.425		.497	
Adjusted R- Square	.170		.238		.334		.410	
R-Square Change	.189		.092		.144		.072	
F Change	10.004***		3.552*		2.710*		10.800**	
Note: $N = 89$ Teachers * $p < .05$ , ** $p < .01$ , *** $p < .001$ X = Coefficients for District ID and School ID variables not released per MET data use agreement. TKA = Teacher Knowledge Assessment Score								

Adding mean TSJOS (Model 4) to Location (model 1), teacher variables (Model 2) and PLATO sub score variables (Model 3) resulted in a statistically significant increase in the proportion of variance in the dependent variable being explained, with a  $R^2$  change of .072,  $F$  change = 10.8,  $p < .01$ . Mean TSJOS significantly predicted student achievement,  $\beta = .336$ ,  $p < .01$ .

### Summary of Findings

Findings provide evidence for the significant relationship between socially just teaching and student achievement. Mean teacher TSJOS score was a significant predictor of student achievement for all students in the section average score regression (Regression #1), as well as in a separate regression after accounting for location, teacher



and student variables (Regression #3), and for SAT9 scores (Regression #11).

While socially just teaching scores were found to be significant predictors for some sub-groups of students, the results should be interpreted with caution as all subgroups had Durbin-Watson scores indicative of the assumption of independence of residuals not being met. With the assumption of independence of residuals not being met, that indicates that the error is not random, and thus linear regression is not well suited for the analysis.

Socially just teaching scores also significantly predicted student achievement, and explained a significant proportion of variation in scores, after accounting for both teacher FtF ELA scores (Regression #12 and 13) and average PLATO subscores (Regressions #14 and 15).

## **Chapter 5 - Discussion**

The overarching objective of this study was to examine the relationship between

socially just teaching practices and student achievement. The educational achievement gap is a social justice issue, as Collopy, Bowman, and Taylor (2012) point out, and thus provides context and relevance to the current study. It is not acceptable to only measure teachers for being successful, as typically measured by statewide achievement results, but our students deserve teachers who get to those successful results through building positive relationships creating a just classroom culture. In other words, this study advocates for the need for teachers to both good and successful, as Fenstermacher & Richardson (2005) describe. The findings add to the current body of knowledge regarding this topic and strengthen evidence (Mitescu et al.; Gutstein, 2007) that increased socially just teaching practices do indeed correlate to increases in student achievement. The discussion points that follow are arranged by major topics from Chapter 2. A discussion around socially just teaching practices and student achievement is followed by interpretation of the relationships between UDL, and then OTL, and student achievement. The findings of this study in the context of teacher VAM and student achievement is then discussed.. Implications to the field and then limitations are discussed prior to ending with concluding thoughts.

### **Socially Just Teaching Practices and Student Achievement**

The primary goal of this study --to add to the current body of research connecting socially just teaching practices to students, was met. The simple linear regression in this study that predicted a teacher's class z-score average on the 2011 state ELA exam using teacher mean TSJOS score found similar results to Mitescu et al. (2011). For example,

Mitiescu et al. found a significant, positive relationship ( $r = .44, p < .05$ ) between socially just teaching and student achievement, with an  $R^2 = .19$ , which indicates 19% of the proportion of variance in student achievement scores was explained by teacher mean TSJOS. The current study also found a significant, positive relationship ( $r = .425, p < .001$ ), with mean TSJOS score accounting for approximately 17% of the variation in student achievement scores (adjusted  $R^2 = .172$ ). With the SAT9 Open Ended assessment as the dependent variable, there was also a significant positive relationship ( $r = .214, p < .001$ ) between socially just teaching practices and student achievement, with an adjusted  $R^2 = .045$ . In addition, the current study found a similar mean TSJOS score for the 107 teachers include in the study ( $M = 2.30, SD = .77$ ) compared to Mitiescu et al. ( $M = 2.23, SD = .66$ ).

The hierarchical regression model that entered location, teacher, and student variables (Table 12/Regression #3) provides further evidence to the relationship between socially just teaching practices and student achievement. After accounting for all of those variables, including a student's previous year state ELA z-score, there was still a significant, positive ( $\beta = .02, p < .05$ ) relationship between socially just teaching practices and student achievement, although the proportion of variance in student achievement scores can be described as minimal at best. These findings, taken together, further strengthen the scholarly narrative around the importance of socially just teaching practices not only to increase equity and opportunity, but as a tool to increase the achievement of all students.

While the current study adds findings in Mitescu et al., it also addresses many of that study's limitations. The assessments in this study consisted of state ELA exams, that were given to all students within a district, and was transformed to a z-score to compare student achievement to their peers across the country. The current study also adds to the current body of research by examining the relationship between ELA scores and socially just achievement, as Mitescu et al. analyzed math scores as the dependent variable. The current study had a larger sample size of teachers with a wide range of teaching experience. The current study also provides evidence that socially just teaching practices have a significant, positive relationship on both traditional statewide assessments used to measure student achievement in addition to assessments, in this case the SAT9, that measure more complex cognitive skills. Separate regression analyses that looked at 4th and 5th grade participants separately on both the 2011 state ELA exam (Table 12) and the SAT9 (Table 13) as outcome variables repeated the result that socially just teaching practice have a positive, significant relationship with student achievement. Interestingly, the difference in magnitude of the correlation between grades were different. When the 2011 state ELA exam was the dependent variable, socially just teaching practices had a  $R = .388$  for 4th graders and  $.490$  for fifth graders, and  $R = .212$  for 4th graders when the SAT9 was the dependent variable and  $.225$  for 5th graders. The overall stronger correlation between teacher mean TSJOS score and student achievement with the outcome variable of 2011 state ELA exam compared to the SAT9 as the outcome variable goes against the findings in Grossman et al. (2014) of the SAT9 being a more sensitive

predictor of PLATO scores than the state ELA exam. A possible reason for this discrepancy in findings is the TSJOS is more aligned with outcomes measured in the state ELA exams, and is not sensitive to higher-level cognitive skills measured in assessments like the SAT9. The findings that socially just teaching practice for 5th graders had a stronger correlation with student achievement (2011 state ELA exam) than for 4th graders, is also interesting. One hypothesis is the TSJOS has a stronger relationship with skills addressed in 5th grade and the strategies associated with the TSJOS has a larger impact on student learning in 5th grade. A future area of research is to conduct a similar analysis to this study focusing on middle and high school classrooms to further investigate any grade level differences in both the significance and magnitude of the relationship between socially just teaching practices and student achievement.

Taken collectively, this study adds to the context and provides further evidence, in terms of standardized dependent variables, a focus on ELA achievement, and a relatively large sample size of teachers and students, to add to the hypothesis that socially just teaching practices are in fact positively correlated with student achievement.

The next set of research questions all included simple linear regressions to analyze the impact of socially just teaching practices on student achievement for the following sub-groups of students: ELL, SWD, free/reduced price lunch, Gifted/Advanced, and traditionally underserved Races (Black, Hispanic and American Indian). While the regression findings indicate interesting findings, all subgroup regressions had Durbin-Watson statistics that indicated the null hypothesis of

independence of residuals should be rejected, and the Durbin-Watson values all indicated evidence of positive autocorrelation. With the assumption of linear regression being violated, the implications of the findings will not be discussed.

The independence of residuals, or error terms, is one of the assumptions of linear regression, and hierarchical multiple linear regression (Pedhazur, 1997). When the assumption is not met, the null hypothesis that the residuals are independent is rejected, and the conclusion that linear regression (or hierarchical multiple linear regression) is not an appropriate analysis must be made. Next steps include continuing to work on validating observation protocols for the subgroups that include strategies to increase student achievement (Jones et al., 2013). In addition, identifying how specific socially just practices are related to, and impact, student achievement is a next step for the research field (Dover, 2009).

### **UDL, OTL, and Student Achievement**

A proposed connection between FfT2 and UDL was presented in the literature review (also see Table 7). While this connection has not been validated, study findings do suggest the alignment of certain FfT subscales with UDL principles, as was suggested by the authors of the FfT and UDL (CAST & The Danielson Group, 2014). The average ELA FfT scores, then, can be said to be aligned to some elements of UDL. Therefore, findings in this study provide initial ideas on the distinct and unique contribution of socially just teaching practices on student achievement above and beyond other observational tools (see Table 21/Regression #12), specifically the FfT and PLATO

observation protocols.

The block of variables that included the eight average ELA subscale scores of the FfT used in the MET study explained approximately 17.9% of the variance in student state ELA rank based z-scores. When teacher mean TSJOS was added to the model (Model 2), however, an additional 7.8% of the variance was explained, which was significant ( $p < .01$ ). In addition, when location variables (district and school) as well as teacher variables (gender, race and knowledge assessment score) were added as the first two models, respectively, mean TSJOS still had a change in the proportion of variance in student achievement of 4.3% above and beyond the FfT scores (see Table 20/Regression #11). These findings suggests that socially just teaching, at the very least, adds prediction value to student achievement, and provides some evidence that the TSJOS provides a unique explanation of student achievement above and beyond the FfT components. While it may be an incomplete bridge, the alignment of the components of the FfT with UDL principles coupled with the finding that socially just teaching practices explain a significant proportion of variance in student achievement after taking FfT scores in consideration, implies that socially just teaching practices may be a distinct concept from UDL. Future research should focus on measuring UDL principles in action, through the validation of an observational protocol, and also dive deeper into the relationship between evidence of UDL principles in the classroom and (i.) student achievement and (ii.) socially just teaching practices.

A similar story can be told when comparing the TSJOS to PLATO scores. Mean

TSJOS explained an additional 12% of the proportion of the variance in student achievement above and beyond what the average scores of the seven PLATO subscales explained (see Table 23/Regression #14). The  $R^2$  change statistic decreased from .120 to .072 when location and teacher variables were entered into the regression analysis as the first two models. However, both regression analyses (Regressions #14 and #15) indicate a significant, positive relationship ( $p < .001$  and  $p < .01$ , respectively) between socially just teaching and student achievement.

This study was designed from an OTL perspective, as it measures the intersection of the student and his/her classroom (Moss et al., 2008; Pullin and Haertel, 2008) by measuring both student output (student achievement) in terms of the context of the class and teacher instruction and measured as socially just teaching practices identified in the TSJOS. A more consistent definition of OTL is warranted to overcome measurement difficulties (Elliot, 2015), and analyzing the significance and magnitude of specific OTL practices for specific groups of students, like ELLs (Aguirre-Munoz and Ambisca, 2010) are two next steps. Future research on the relationship between OTL best practices, once defined, and socially just teaching practices, will provide further information on the distinction between the concepts of socially just teaching and OTL.

In summary, the evidence from this study suggests that socially just teaching practices, as measured by the TSJOS, does indeed measure a unique construct of socially just teaching, above and beyond standardized classroom observation that measure both effective teaching (FFT) as well as effective language art teaching (PLATO). The study



also provides evidence to support the claim that socially just teaching practices have a positive and significant correlation with student achievement.

### **Teacher VAMs and Student Achievement**

As Grossman (2014) found, the relationship between teacher Protocol for Language Arts Teaching Observation (PLATO) and VAM scores varies depending on the type of assessment used to measure student achievement. The current study adds to the field by suggesting that the type of assessment also matters when measuring socially just teaching practices. Mean TSJOS was a significant, positive predictor of student achievement as defined as both SAT9 z-score scaled score and 2011 state ELA rank based z-score at the  $p < .001$  level. Mean TSJOS, however, explained only 4.5% of the variance in student SAT9 scores compared to 17.2% of the variance in state ELA exam rank based z-scores. This result goes against the hypothesis that the TSJOS would be a stronger predictor of higher-level assessments, like the SAT9, because of the emphasis of the TSJOS items on involving students in higher-level thinking skills, like critiquing procedures and encouraging divergent modes of thinking.

### **Implications**

This study contributes to the current body of research on the relationship between socially just teaching practices and student achievement. It is recommended that the TSJOS can be used as a tool not just for teacher preparation programs with a social justice mission/focus as Mitescu et al. suggest, but as a means to support good and successful teaching across sectors, including public, private and charter schools. The idea

of social justice as transparent communication and commitment to high achievement for all (Perry, Steele and Hilliard, 2003) aligns with the idea of utilizing the TSJOS as a feedback tool for teachers across organizations that share the common theme of socially just teaching. An important distinction between teaching for social justice and socially just teaching is warranted. Teaching for social justice is an idea that describes the act of purposely and explicitly teaching students about current inequities, and providing students with strategies and means to increase their equity and opportunities. Socially just teaching practices are practices that lead to a socially just learning environment, but does not necessarily require the teacher to directly teach about current inequities and ways to maximize equity and equality. This study is not advocating or measuring the impact or relationship between teaching for social justice and student achievement, but is rather focused on the teaching practices that are believed to support teachers in providing an equitable and socially just education to their students.

The wide range of years of teaching experience in the sample of teachers analyzed in this study adds to the notion that following indicators spelled out in the TSJOS can be a useful tool for teachers on various points of their professional careers can benefit from. Reflecting on practices included in the TSJOS would be a beneficial exercise for teachers regardless of experience, and the TSJOS could be used in a variety of settings, from school-wide learning walks or instructional rounds, to individual teacher coaching sessions.

The current findings do not suggest using the TSJOS for high stakes hiring

decisions as Winters and Cowen (2013) describe some states using VAMs. It is recommended that educational organizations leverage the TSJOS to provide coaching and feedback to teachers as a way to reflect and improve their practice in hopes of improving good (socially just) teaching practices and successful (student achievement) teaching practices.

### **Limitations**

There are a number of limitations and related next steps for future research. A major limitation of the study is the regression analyses that attempted to predict student achievement from socially just teaching practices for subgroups of interest not meeting the linear regression assumption of normality of residuals. Future research focusing on the relationship between socially just teaching practices and student achievement for traditionally marginalized students, including ELLs, SWDs, students who qualify for free/reduced price lunch students who identify with traditionally underserved races will add to the research community's understanding of the different relationships between socially just teaching and student achievement for these groups.

The current study also did not include teacher variables that could be contributing to some of the variability in student achievement scores. More specifically, years of teaching experience and teacher education (Master's degree or not) were two variables not included in the study due to a high percentage of missing data for these variables. Years of teaching matters as a study focusing on socially just teaching and student achievement advocated for the use of the TSJOS for novice teachers (Mitescu et al.). The

decision was made to include teachers regardless of years of teaching experience for two reasons. The first reason was to have a sample for this secondary analysis that would be sensitive enough to pick up on differences in the data based on a power analysis. The second reason was to assess the predictive power of socially just teaching practices in terms of student achievement for all teachers so a claim could be made around possibly using the TSJOS in educational setting for a feedback tool for teachers at various points in their professional journey. Teachers' education have been included in VAM models (Braun, Jenkins, Grigg, and Tirre, 2006; Bill and Melinda Gates Foundation, 2013) trying to identify effective teachers, as defined by student achievement.

Garrett and Steinberg (2015) determined the extent of student non-random assignment in the MET database made it difficult to infer causality between teacher VAM scores and student achievement scores. While many students were randomly assigned to the teacher they actually had for the year for the data being analyzed, it is important to note the impact of non-random assignment resulting in the inability to infer causality from the analysis, as Garrett and Steinberg found. This study does permit comment and adds to the body of knowledge on the relationship between socially just teaching and student achievement, but inferring causality is beyond the scope of this project.

Future directions for research based off of this study include expanding the settings to describe the relationship between socially just teaching and student achievement. Future research can focus on a wider range of grades, with a focus on

middle and secondary school, for example. Analyzing the differential impact of specific items, or groups of items on the TSJOS on student achievement would also add a nuanced understanding of the relationship of specific strategies and practices that are socially just on student achievement is another area for possible future inquiry.

As mentioned previously, research focusing on specific subgroups of students in terms of socially just teaching practices is important. Although the sub-group analyses indicated a violation of the assumption of independence of errors, the results suggested that socially just teaching practices had a stronger relationship with student achievement for the majority group (non-ELL, non-gifted, etc.) across sub-groups. A deeper analysis into the impact and relationship of socially just teaching practices with student achievement that for sub-groups is warranted. Part of this future research would also include a comparison of the findings for each subgroup (ELL students, for example) verses the reference group (non-ELL students, for example). Without an understanding of the how socially just teaching practices are related to student achievement for traditionally marginalized groups, it hinders efforts to strengthen the evidence that points to socially just teaching as a means not just for equity but for increased student achievement for traditionally underserved and marginalized students.

In addition, the composition of classrooms in terms of student demographics is another area that could make for interesting and informative research. In other words, research could focus on analyzing the effect of having a classroom with a relatively high percentage of students with disabilities, or students who are Gifted/Advanced, on

individuals with the same characteristics. For example, the a future research design could analyze the relationship between socially just teaching practices and student achievement for individual students in sub-groups of interest, and include a predictor variable of the percentage of students in that individuals classroom who had the same characteristic as part of the regression model and analysis.

### **Conclusion**

The current achievement gap in the Unites States is a social justice issue (Collopy, Bowman, and Taylor, 2012). Therefore, social justice can be viewed as both a set of opportunities, or providing an opportunity to learn for all students, and as having increased achievement being one its primary goals. The distinction of good and successful teaching (Fenstermacher & Richardson, 2005) is important. Social justice demands that teachers are not just successful, or effective at increasing achievement results, but obtain increased achievement through moral and just teaching practices. The present study adds to the existing evidence that socially just teaching practices, or good teaching, is indeed related to successful teaching, on two different measurements of student achievement. Socially just teaching was found to explain a unique proportion of variation in student achievement scores above and beyond two commonly used teacher observation scales (the FfT and PLATO). Socially just teaching practices is one of many frameworks that can be employed to bring an educational organization closer to its collective mission of social justice and equity. This mission can be realized by increasing student achievement, so that all students are prepared to not only realize their dreams, but

have the agency to expand and realize their dreams.

#### References

Aguirre-Munoz, Z., & Amabisca, A. (2010). Defining opportunity to learn for English

- Language Learners: Linguistic and cultural dimensions of ELLs' instructional contexts. *Journal of Education for Students Placed at Risk*, 15(3), 259-278.
- Albano, A.D., & Rodriguez, M.C. (2013). Examining differential math performance by gender and opportunity to learn. *Educational and psychological measurement*, 73(5), 836-856.
- Alliance for Excellence in Education (2010). *High school dropouts in America*. Fact Sheet. Washington, DC: Author. Retrieved from [all4ed.org/wp-content/uploads/HighSchoolDropouts.pdf](http://all4ed.org/wp-content/uploads/HighSchoolDropouts.pdf)
- Artiles, A.J., Bal, A., & Thorius, K.A.K. (2008). Back to the future: A critique of response to intervention's social justice views. *Theory into practice*, 49, 250-257.
- Baglieri, S., & Knopf, J.H. (2004). Normalizing difference in inclusive teaching. *Journal of learning disabilities*, 37(6), 525-529.
- Bill and Melinda Gates Foundation (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study*. Seattle, WA: Author.
- Bill and Melinda Gates Foundation (2010a). *Learning about teaching: Initial findings from the Measures of Effective Teaching Project*. Seattle, WA: Author.
- Bill and Melinda Gates Foundation (2010b). *Danielson's framework for teaching for classroom observations*. Seattle, WA: Author.
- Braun, H. (2005). Using student progress to evaluate teachers: A primer on value-added models. Princeton, NJ: Educational Testing Service.



- Braun, H., Jenkins, F., Grigg, W., & Tirre, W. (2006). *A closer look at charter schools using hierarchical linear modeling*. Washington, DC: National Center for Education Statistics.
- Boote, D.N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational Researcher*, 34(6), 3-15.
- Brooks, K., Adams, S.R., & Morita-Mullaney, T. (2010). Creating inclusive learning communities for ELL students: Transforming principals' perspectives. *Theory into practice*, 49, 145-151.
- Callahan, R.M. (2005). Tracking and high school English Language Learners: Limiting opportunity to learn. *American Education Research Journal*, 42(2), 305-328.
- CAST & The Danielson Group (2014). *Crosswalk between Universal Design for Learning (UDL) and the Danielson Framework for Teaching (FfT)*. Wakefield, MA: National Center on Universal Design for Learning.
- CAST (2011). *UDL Guidelines – Educator's Worksheet Version 2.0*. Retrieved from <http://www.udlcenter.org/aboutudl/udlguidelines/downloads> on June 20, 2015. Wakefield, MA: National Center on Universal Design for Learning.
- CAST (2007). *Summary of 2007 national summit on universal design for learning working groups*. Wakefield, MA: Author.
- Castro-Villarreal, F., & Nichols, S.L. (2016). Intersections of accountability and special education: The social justice implications of policy and practice. *Teachers college*

- record*, 118, 1-24.
- Cawthon, S.W., Beretvas, A.D., & Lockhart, L.L. (2012). Factor structure of opportunity to learn for students with and without disabilities. *Educational Policy Analysis Archives*, 20(41), 1-26.
- Collopy, R., Bowman, C., & Taylor, D.A. (2012). The educational achievement gap as social justice issue for teacher educators. *Catholic Education*, 16(1), 4-25.
- Cooper, R., & Liou, D.D. (2007). The structure and culture of information pathways: Rethinking opportunity to learn in urban high schools during ninth grade transition. *The High School Journal*, 91(1), 43-56.
- Danielson, C. (2007). *Enhancing Professional Practice: A Framework for Teaching* (2nd Edition). Alexandria, VA: ASCD.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, R. (2012). Evaluating teacher evaluation. *The Phi Delta Kappan*, 93(6), 8-15.
- Dellinger, A.B. (2005). Validity and the review of literature. *Research in Schools*, 12(2), 41-54.
- Dougherty, K.J. (1995). Opportunity-to-learn standards: A sociological critique. *Sociology of Education*, 69(40), 40-65.
- Dover, A.G. (2009). Teaching for social justice and K-12 student outcomes: A conceptual framework and research review. *Equity and excellence in education*, 42(4), 506-524.
- Echevarria, J., Vogt, M., & Short, D.J. (2013). *Making Content Comprehensible for English Learners: The SIOP Model* (4th ed.). Boston, MA: Pearson.

- Elliott, S.N. (2015). Measuring opportunity to learn and achievement growth: Key research issues with implications for the effective education of all students. *Remedial and Special Education, 36*(1), 58-64.
- Fenstermacher, G.D., & Richardson, V. (2005). On making determinations of quality in teaching. *Teachers College Record, 107*(1), 186-213.
- Freire, P. (1993). *Pedagogy of the Oppressed*. New York: Continuum.
- Fry, R. (2007). *How Far Behind are English Language Learners?* Washington, DC: Pew Hispanic Center. Ganley, P. & Ralabate, P. (2013). *UDL Implementation: A Tale Of Four Districts*. National Center on Universal Design for Learning. Retrieved May 18, 2015 from <http://www.udlcenter.org/implementation/fourdistricts/bartholomew> .
- Garrett, R., & Steinberg, M.P. (2015). Examining teacher effectiveness using classroom observation scores: Evidence from the randomization of teachers to students. *Education evaluation and policy analysis, 37*(2), 224-242.
- Goldhaber, D.D., Goldschmidt, P., & Tseng, F. (2013). Teacher value-added at the high-school level: Different models, different answers? *Educational Evaluation and Policy Analysis, 35*(2), 220-236.
- Grant, C. A. & Agosto, V. (2008). Teacher Capacity and Social Justice in Teacher Education. In M. Cochran-Smith, S. Feiman Nemser, J. McIntyre, & K. Demers (Eds.), *Handbook of research on teacher education: Enduring questions in changing contexts* (pp. 175-200). Mahwah, NJ: Lawrence Erlbaum Publishers.

- Grossman, P., Cohen, J., Ronfeldt, M., & Brown, L. (2014). The test matters: The relationship between classroom observation scores and teacher value added on multiple types of assessments. *Educational Researcher*, 43(6), 293-303.
- Gutstein, E. (2007). "And that's just how it starts": Teaching mathematics and developing student agency. *Teachers college record*, 109(2), 420-448.
- Hanushek, E.A. (2002). Teacher quality. In L.T. Izumi & M.E. Williamson (Eds.). Teacher quality (pp.1-12). Stanford, CA: Hoover Press.
- Harris, D.N., Ingle, W.K., & Rutledge, S.A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added models. *American Educational Research Journal*, 51(1), 73-112.
- Heafner, T.L., & Fitchett, P.G. (2015). An opportunity to learn US History: What NAEP data suggest regarding the opportunity gap. *The High School Journal*, 98(3), 226-249.
- Hemphill, F.C., Vanneman, A., & Rahman, T. (2011). *Achievement gaps: How Hispanic and white students in public schools perform in mathematics and reading on the national assessment of educational progress: A statistical report*. Washington, DC: National Center for Education Statistics.
- Higher Education Opportunity Act of 2008 (Public Law 110-315).
- Hill, H.C., Kapitula, L., & Umland, K. (2011). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal*,

- 48(3), 794-831.
- Hitchcock, Meyer, A., Rose, D., & Jackson, R. (2002). Providing new access to the general curriculum: Universal design for learning. *Teaching Exceptional Children, 35*(2), 8-17.
- Inter-University Consortium for Political and Social Science Research (n.d.). *Measures of effective teaching 1: Study information - Video information file codebook*. Ann Arbor, MI: Author.
- Kauffman, J.M., Anastasiou, D., Badar, J., Travers, J.C., & Wiley, A.L. (2016). Inclusive education moving forward. *Advances in special education, 32*, 153-178.
- LaFontaine, D., Baye, A., Vieluf, S., & Monseur, C. (2015). Equity in opportunity-to-learn and achievement in reading: A secondary analysis of PISA 2009 data. *Studies in Educational Evaluation, 47*, 1-11.
- Lefgren, L., & Sims, D. (2012). Using subject test scores efficiently to predict teacher value-added. *Educational Evaluation and Policy Analysis, 34*(1), 109-121.
- Jaafar, S.B. (2006). An alternative approach to measuring opportunity to learn in high School classes. *The Alberta Journal of Educational Research, 52*(2), 107-126.
- Joe, J.N., McClellan, C.A., & Holtzman, S.L. (2014). Scoring design decisions: Reliability and the length and focus of classroom observations. In T.J. Kane, K.A. Kerr & R.C. Pianta (Eds.). *Designing Teacher Evaluation Systems* (pp. 415-443). San Francisco, CA: Jossey-Bass.
- Jimenez, Terese C., Graf, Victoria, L., & Rose, Ernest, (2007). Gaining access to general:

- The promise of universal design for learning. *Issues in Teacher Education*, 16(2), p. 41-54.
- Karl, A.T., Yang, Y., & Lohr, S.L. (2013). A correlated random effects model for nonignorable missing data in value-added assessment of teacher effects. *Journal of Educational and Behavioral Statistics*, 38(6), 577-603.
- Kauffman, J.M., & Badar, J. (2014). Instruction, not inclusion, should be the central issue in special education: An alternative view from the USA. *Journal of international special needs education*, 17(1), 13-20.
- Kavale, K. A., & Forness, S. R. (2000). History, rhetoric, and reality: Analysis of the inclusion debate. *Remedial and Special Education*, 21(5), 279-297.
- Kilbourn, B. (2006). The qualitative doctoral dissertation proposal. *Teachers College Record*, 108(4), 529-576.
- Kurz, A., Elliott, S.N., Wehby, J.H., & Smithson, J.L. (2010). Alignment of the intended, planned and enacted curriculum in general and special education and its relation to student achievement. *The Journal of Special Education*, 44(3), 131-145.
- Kurz, A., Talapatra, D., & Roach, T. (2012). Meeting the curricular challenges of inclusive assessment: The role of alignment, opportunity to learn, and student engagement. *International Journal of Disability, Development and Education*, 59(1), 37-52.
- Ladson-Billings, G. (2006). From the achievement gap to education debt: understanding achievement in U.S. schools. *Educational researcher*, 35(7), 3-12.

- Ladson-Billings, G. (2009). *The Dreamkeepers: Successful Teachers of African American Children*. San Francisco, CA: Jossey-Bass.
- Lather, P. (1999). To be of use: The work of reviewing. *Review of Educational Research*, 69(1), 2-7.
- Maranto, R. & Ritter, G. (2014). Why KIPP is not corporate: KIPP and social justice. *Journal of School Choice*, 8(2), 237-257.
- Mitescu, E., Cochran-Smith, M., Pedulla, J.J., Cannady, M., & Jong, C. (2011). Measuring practices of teaching for social justice in elementary mathematics classrooms. *Education Research Quarterly*, 34(3), 15-39.
- Mihaly, K., McCaffrey, D.F., Staiger, D.O., & Lockwood, J.R. (2013). *A Composite Estimator of Effective Teaching*. Measures of Effective Teaching Project, Bill and Melinda Gates Foundation.
- Mo, Y., Singh, K., & Chang, M. (2013). Opportunity to learn and student engagement: a HLM study on eighth grade science achievement. *Educational Research for Policy and Practice*, 12, 3-19.
- Moss, P.A., Pullin, D.C., Gee, J.P., Haertel, E.H., and Young, L.J. (2008). *Assessment, Equity and Opportunity to Learn*. New York: Cambridge University Press.
- National Center on Universal Design for Learning (2012). *UDL guidelines - version 2.0 - research evidence*. Retrieved from <http://www.udlcenter.org/research/researchevidence>.
- Nichols, S.L., Glass, G.V, & Berliner, D.C. (2012). High-stakes testing and student

- Achievement: Updated analyses with NAEP data. *Education Policy Analysis Archives*, 20 (20), 1-30.
- Novack, M. (2000). Defining social justice. *First things*, 108, 11-13.
- Papay, J.P. (2011). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *American Educational Research Journal*, 48(1), 163-193.
- Paufler, N.A., & Amrein-Beardsley, A. (2014). The random assignment of students into elementary classrooms: Implications for value-added analyses and interpretations. *American Educational Research Journal*, 51(2), 328-362.
- Pedhazur, E.J. (1997). *Multiple regression in behavioral research* (3rd ed.). USA: Thomson Learning.
- Perry, T., Steele, C., & Hilliard III, A. (2003). *Young, Gifted and Black: Promoting High Achievement Among African-American Students*. Boston, MA: Beacon Press.
- Polikoff, M.S. (2014). Does the test matter? Evaluating teachers when tests differ in their sensitivity to instruction. In T.J. Kane, K.A. Kerr & R.C. Pianta (Eds.). *Designing Teacher Evaluation Systems* (pp. 278-302). San Francisco, CA: Jossey-Bass.
- Polikoff, M.S., & Porter, A.C. (2014). Instructional alignment as a measure of teaching quality. *Educational Evaluation and Policy Analysis*, 36(4), 399-416.
- Poplin, M. S. (1988). The reductionist fallacy in learning disabilities: Replicating the past by reducing the present. *Journal of Learning Disabilities*, 71(7), 389-400.
- Rao, K., Ok, M.W., & Bryant, B.R. (2014). A review of research on Universal Design



- educational models. *Remedial and Special Education*, 35(3), 153-166.
- Rose, D., & Meyer, A. (n.d.) *The future is in the margin: The role of technology and disability in education reform*. CAST.
- Pullin, D., and Haertel, E.H. (2008). Assessment through the lens of “Opportunity to Learn.” In Moss, P.A., Pullin, D.C., Gee, J.P., Haertel, E.H., and Young, L.J. (Eds.). *Assessment, Equity and Opportunity to Learn* (pp. 17-41). New York: Cambridge University Press.
- Reeves, E.B. (2012). The effects of opportunity to learn, family socioeconomic status, and friends on rural math achievement gap in high school. *American Behavioral Scientists*, 56(7), 887-907.
- Reeves, C., Carnoy, M., & Addy, N. (2013). Comparing opportunity to learn and student achievement gains in african primary schools: A new approach. *International Journal of Educational Development* 33, 426-435.
- Reichow, B. & Volkmar, F. R. (2010). Social skills interventions for individuals with Autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders*, 40, 149-166.
- Rose, D. H. & Meyer, A. (2002) *Teaching Every Student in the Digital Age: Universal Design for Learning*. Alexandria, VA: Association for Supervision and Curriculum Development. Online version:  
<http://www.cast.org/teachingeverystudent/>
- Scherff, L., & Piazza, C. (2008). Why now, more than ever, we need to talk about

- opportunity to learn. *Journal of Adolescent and Adult Literacy*, 52(4), 343-352.
- Schochet, P.Z., & Chiang, H.S. (2013). What are error rates for classifying teacher and school performance using value-added models? *Journal of Educational and Behavioral Statistics*, 38(2), 142-171.
- Schulte, A.C., Stevens, J.J., Elliott, S.N., Tindall, G., & Nese, J.F.T. (2016). Achievement gaps for students with disabilities; Stable, widening, or narrowing on a state-wide reading comprehension test? *Journal of Educational Psychology*, 108(7), 925-942.
- Shavelson, R.J. (1996). *Statistical Reasoning for the Behavioral Sciences* (3rd ed.). Boston, MA: Allyn and Bacon.
- Shealey, M.W., Sparks, C.W., & Thomas, N. (2012). Defining leadership in charter schools: Examining the intersection of social justice and special education. *Journal of special education leadership*, 25(1), 15-24.
- Skinner, C. (2012). Location and place: Does ELL classroom location reflect US claims of social justice and democracy? *Journal of philosophy and history of education*, 62(1), 81-94.
- Span, C., & Rivers, I.D. (2012). Reassessing the achievement gap: An intergenerational comparison of African American student achievement before and after compensatory education and the Elementary and Secondary Education Act. *Teachers College Record*, 114 (6), 1-17.
- Stotsky (2009, November 13). Who needs mathematicians for math anyway? *City*

*Journal*. Retrieved January 22 2016 from

<http://www.city-journal.org/2009/eon1113ss.html>.

Swanson, L. & Hoskyn, M. (1998). A synthesis of experimental intervention literature for students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research*, 68, 271-321.

Tarr, J.E., Grouws, D.A., Chavez, O., & Soria, V.M. (2013). The effects of content organization and curriculum implementation on students' mathematics learning in second-year high school courses. *Journal for Research in Mathematics*, 44(4), 683-729.

Templeton, N.R. (2011). *Understanding social justice: Improving the academic achievement of African American students*. Ypsilanti, MI: National Council of Professors of Education Administration.

Theoharis, G., & O'Toole, J. (2011). Leading inclusive ELL: Social justice leadership for English Language Learners. *Education administration quarterly*, 47(1), 646-688.

Tienken, C.H., & Achilles, C.M. (2009). Relationship between class size and students' opportunity to learn writing in middle school. *Research in the Schools*, 16(1), 13-24.

Tornroos, J. (2005). Mathematics textbooks, opportunity to learn, and student achievement. *Studies in Educational Evaluation*, 31, 315-327.

United States Department of Education (2009). *Race to the Top Program Executive Summary*. Washington, DC: U.S. Department of Education.

- Wade, R.C. (2007). *Social Studies for Social Justice: Teaching Strategies for the Elementary Classroom*. New York: Teachers College Press.
- Vanneman, A., Hamilton, L., Anderson, J.B., & Rahman, T. (2009). *Achievement gaps: How black and white students in public schools perform in mathematics and reading on the national assessment of educational progress: A statistical report*. Washington, DC: National Center for Education Statistics.
- Wang, A.H. (2010). Optimizing early mathematics experiences for children from low-income families: A study on opportunity to learn mathematics. *Early Childhood Education Journal*, 37, 295-302.
- Wang, J. (1998). Opportunity to learn: The impacts and policy implications. *Educational Evaluation and Policy Analysis*, 20(3), 137-156.
- Wang, J., & Goldschmidt, P. (2003). Importance of middle school mathematics on high school mathematics achievement. *The Journal of Educational Research*, 97(1), 3-19.
- Wang, J., & Goldschmidt, P. (1999). Opportunity to learn, language proficiency, and immigrant status effects on mathematics achievement. *The Journal of Educational Research*, 93(2), 101-111.
- Weisman, E.M., Flores, S.Y., & Valenciana, C. (2008). Coming to voice: Preparing bilingual-bicultural teachers for social justice. *Equity and excellence in education*, 41(3), 357-371.
- Wiley, D.E., & Yoon, B. (1995). Teacher reports on opportunity to learn: Analysis of the

- 1993 California learning assessment system (CLAS). *Education Evaluation and Policy Analysis*, 17(3), 355-370.
- Winters, M.A., & Cowen, J. M. (2013). Who would stay, who would be dismissed? An empirical consideration of value-added teacher retention policies. *Educational Researcher*, 42(6), 330-337.
- Zeichner, K.M. & Conklin, H.G. (2005). Teacher education programs. In M. Cochran-Smith & K.M. Zeichner (Eds.), *Studying Teacher Education: The Report of the AERA Panel on Research and Teacher Education* (pp. 645-735). Mahwah, NJ: Lawrence Erlbaum Associates.