Spring 2016

Empowering School Leadership across Nations: A Comparative Study of Differences Between Leadership Practices and Their Relationships with Empowerment Amongst Administrators and Teachers in Eight Large Urban Secondary School Districts in the United States and Canada

Anya Niazov

DePaul University

College of Education

EMPOWERING SCHOOL LEADERSHIP ACROSS NATIONS: A COMPARATIVE STUDY OF DIFFERENCES BETWEEN LEADERSHIP PRACTICES AND THEIR RELATIONSHIPS WITH EMPOWERMENT AMONGST ADMINISTRATORS AND TEACHERS IN EIGHT LARGE URBAN SECONDARY SCHOOL DISTRICTS IN THE UNITED STATES AND CANADA

A Dissertation in Education

with a Concentration in Educational Leadership

by

Anya Niazov

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Education

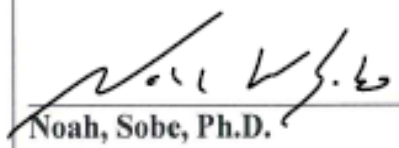
June 2016

The dissertation of Anya Niazov is hereby approved.

6/1/2016

Joseph Gardner, Ph.D.
Associate Professor
DePaul University
Chair of Committee

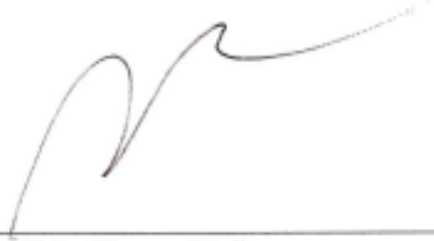
Date



6/1/2016

Noah, Sobe, Ph.D.
Professor
Loyola University
Committee Member

Date



6/1/16

An Chih Cheng, Ph.D.
Assistant Professor
DePaul University
Committee Member

Date

ABSTRACT

U.S. schools have been shown to underperform other industrialized nations on tests of academic achievement. Canada is a higher-performing nation on international assessments and has a similar educational system to the United States. This study aims to determine if there are differences between urban U.S. and Canadian secondary schools and specifically addresses school leadership in these schools. The study employs transformational leadership theory to assess cross-district and cross-national differences in leadership behaviors. In this quantitative study, I characterize differences in administrator and teacher leadership behaviors, teacher satisfaction and empowerment, and their correlations within North American schools. Survey data were collected from over 500 secondary teachers and administrators and results were analyzed using inferential statistical methodologies.

This study demonstrates that there are significant differences between teacher leadership and empowerment both across districts and cross-nationally, but that there are not cross-national differences across administrative leadership. Administrator leadership did, however, have a substantial impact on teachers at the level of the school. Self-reported administrative leadership behaviors not only correlated with teachers' likelihood to demonstrate similar classroom leadership behaviors, but also were associated with teachers having higher levels of self-efficacy in areas such as "autonomy" and "status".

There were many significant differences in teacher leadership and teacher empowerment between two large districts (one in the US and one in Canada) and cross-nationally. While some of these differences were related to administrator leadership, this effect did not fully explain the differences among the teachers. This work builds on existing scholarship in the areas of

educational leadership and international education and extends this work to assess the impact of leadership on teacher empowerment in a cross-district and cross-national context. My study also importantly lays the foundation for ongoing scholarship that can further illuminate the causes for differences and provides a critical knowledge base to help researchers and practitioners determine how impacting leadership behaviors can improve teacher empowerment and other educational outcomes in schools both in the U.S. and internationally.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEDICATION

My kind, warm, loving and incredibly supportive grandmother, whom I reached out to regarding any difficult situation, was exhilarated that I followed her footsteps by becoming a school teacher, a mentor, and an advisor. As I pursued my love for teaching and learning by embarking on my doctoral journey in education, she continued to be incredibly proud of me and always gave me the most pertinent and meaningful advice. Unfortunately, my babooshka whose personality embodied the very essence of selfless care, depth of a meaningful life, and infinite love, was not able to witness the finale of my dissertation process, because at the wise age of eighty eight, Enya (Reznik) Niazova, departed this world in January of 2011. I love you, your spirit is always with me, and I know you are proud of me. It is my greatest privilege to dedicate my meaningful doctoral dissertation in your honor. You have been an inspiration to me and your presence will continue guiding me throughout life.

ACKNOWLEDGEMENTS

The deepest, greatest, heartfelt, and most significant thank you is bestowed to my partner, my friend, my confidant, my husband, the love of my life, Benjamin. Thank you for your unconditional love, your loyal friendship and your unequivocal support throughout this arduous dissertation process. Thank you for being my professional editor when I could no longer see or understand what was appearing in front of me after spending eighteen hours a day on researching, writing and editing for the past several years. Thank you for being my personal masseur when I could no longer feel my hands, brain or shoulders after not moving for days at a time. Thank you for encouraging, believing in and calming me throughout the dissertation years when I have not slept for days, I was not myself and I felt that life no longer had meaning. Thank you for working additional day and night shifts in order to send me to an international conference related to my dissertation and to help pay for student loans. Thank you for thinking and processing life differently, I hope you know how much your love and support mean to me, because without it, I would not be alive today.

I have always and will continue to thank my parents for instilling the love and importance of education. I am grateful to have highly educated parents and always strived to make my mom (a valedictorian all throughout her undergraduate and graduate studies) and my dad (who to this day, continues to elevate his understanding and intellect by constantly learning new things at any age) proud and I hope you are! I am humbled by the sacrifices (including your careers) that you had to make in order to relocate (without a single word of English) all of our lives to America in hope of a better life for all of us. Thank you for your unique love and care.

I hope I can be certain that Sobachka fully understands what his love, support and presence means to me. Throughout the past seven challenging years, my brain thought about the dissertation every waking and sleeping minute of each day, except when I interacted with this significant being. Thank you for not holding grudges and being a positive example in my life. Without consciously intending, you make me smile every day and I will always treasure our relationship.

A meaningful thank you belongs to A.C. for sincerely listening and getting to know me as a sincere, reflective, critical, highly intellectual, passionate student of self and life. Thank you for being able to calmly redirect my focus onto my strengths and somehow use just the right words to suggest an effective direction for me in order to encourage me to hang in there and try to remain sane.

I am grateful to my fellow doctoral colleague, Olga Buchko, who offered her long distance shoulder to cry on and ears to listen as well as moral support as I progressed through each of the individual obstacles along my dissertation process. Thank you for providing guidance and motivation as we each dealt with the trials, traumas, and tribulations of our individual dissertation processes.

A collegial thank you goes to Dr. Suzanne Velasquez-Sheehy who cared about me and continued to check up on me throughout the dissertation years while offering inspirational suggestions to help me move along the process. As my colleague, as a teacher, as an administrator, and as someone who has been in the educational sphere for over twenty five years, I thank you for supporting me and believing in the significance and importance of my research.

I chose Dr. Joseph Gardner to be my Chair because he was the only professor, out of all the classes that I have taken during my doctoral program, who was interested in reflecting and improving his teaching by asking his students to evaluate his efficacy half way through the quarter. I was tremendously impressed at this professor's willingness to work on his teaching strategies in order to benefit the learning of his doctoral students. Thank you for your critical reading of the manuscript at every stage of the process, challenging me, and providing important literature suggestions.

My sincere gratitude goes out to a comparative and international education committee member, Dr. Noah Sobe from Loyola University. Despite your various responsibilities and leadership roles within your university and throughout professional organizations, you met with me (even though I was not your student), volunteered your time, offered insightful theoretical suggestions, and fundamental steps that helped me progress from conceptual to practical stages of the process.

Thank you Dr. An Chih Cheng, my committee member, for providing statistical suggestions in my methodology section. Your teaching throughout my two quantitative classes and affirmation of my thoughtful statistical understanding and abilities effectively contributed to my quantitative focus throughout my methodological approach, testing, analysis and interpretation.

An important thank you goes to Dr. David Degras, a professor of statistics from DePaul University, who responded to a doctoral student in need of advanced statistical consultation from the College of Education. Thank you for offering your time and statistical suggestions by helping me explore the nuances and implications of various statistical testing methods.

A statistical consultant from Northwestern University who chose to remain anonymous provided numerous hours of inspecting my work by providing feedback towards quantitative SPSS tests, analysis and interpretations. Thank you for viewing your help to a doctoral student in need of second pair of statistical eyes performing a meaningful educational research as a selfless honor.

Data analysis would not have taken place without the voluntary anonymous participants, who comprised the data, hence I would like to extend a very special thank you to almost five hundred teachers and administrators who dedicated their time, without any direct incentive (but maybe an understanding that my research is meaningful and a hope or a belief, that their teachers and school administrators, as well as other schools, might benefit from study findings), to honestly complete the online survey.

EPIGRAPH

The issue is not just the standards that are written on paper. It's the entire teaching and learning that is created in high-achieving countries like Finland and Singapore, Australia, Canada, Hong Kong and others, that is very different from what we have in the United States.
Linda Darling Hammond

Bureaucratic solutions to problems of practice will always fail because effective teaching is not routine, students are not passive, and questions of practice are not simple, predictable, or standardized. Consequently, instructional decisions cannot be formulated on high then packaged and handed down to teachers.
Linda Darling Hammond

If your actions inspire others to dream more, learn more, do more and become more, you are a leader.
John Quincy Adams

The mediocre teacher tells. The good teacher explains. The superior teacher demonstrates. The great teacher inspires.
William Arthur Ward

A. INTRODUCTION

The United States holds a position of international prestige and hegemony (Cross, 2010). Given this political and economic status, many people from around the world see emigration to the U.S. as a source of better opportunities, both financial and educational (Fowler, 2013). International families from many countries regard the possibility of moving to the U.S. with hope, optimism and with the dream of creating an American life for their children and future generations. My immediate and extended family escaped oppression, persecution, and experienced the great difficulties of leaving their lives in the former Soviet Union to come to the United States. Since coming here, I have had the opportunity to meet and work with many people from all over the world who have experienced similar emigration and transition. While most of us are living better lives in this country, from social and economic perspectives, international test data suggest that the children who attend public schools in the United States are academically behind the students of many of the countries they came from.

If this country is a leader in so many other ways, why do American schools, in general, lag behind other nations' as evidenced by the results of international tests?

Study Topic And Rationale

International Comparisons In Education: A Human Capital Perspective

The modern world is becoming more connected and interdependent. The goals, curricula, and the structure of school systems across the globe are becoming more like each other in order to meet the increasingly common educational needs of modern society (Fowler, 2013; Baker & LeTendre, 2005). Students from all countries either are or will soon be competing not

only with those from their own localities, but internationally (OECD, 2010). Given this new connected and information-driven world, the jobs of tomorrow will go to the most educated students from around the world (Paine & Schleicher, 2011). Therefore, the requirement for high quality education has become essential for students' and nations' long-term success. A nation's education system plays a major role in defining whether students will be prepared for the global world (Tucker, 2011).

In this context of cross-national education, it is increasingly important to establish the relative accomplishment of national education systems in order to assess their current status and determine how they can improve. While international comparative educational tests such as the Program for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS) have limitations, they do serve as an important role as a barometer for assessing national areas of educational strength and weakness. In this dissertation, I focus on this view of the tests as an instructive guide to the relative success and efficacy of various national school systems rather than attempting to use the data to make an economic argument. These international tests have consistently shown that U.S. students perform at average to below average levels compared to other nations (OECD 2012; Tucker 2011). By comparison, nations such as Japan, Singapore, Finland and Canada are top performers on these evaluations (Tucker, 2011). These nations have structured educational reforms to focus on limited, clear learning goals that emphasize depth over breadth, fair resource distribution, supports for teachers, and re-organizing schools to promote learning (Darling-Hammond, 2010). Furthermore, these countries have designed educational frameworks based on research from abroad and then tailored them to

implement their own curricula, cater to their students, and provide a comparative advantage (Tucker, 2011; Fowler, 2013; Darling-Hammond, 2010).

Many argue that the U.S. educational system is still adequate because the best U.S. students are still competitive with the best from other nations (Bracey 2009; Baker 2007). However, educational inequity, particularly differences between rich and poor students, is among the biggest shortcomings of the U.S. educational system compared to other nations. U.S. urban schools where poverty is common are among the most severely affected, and students from these schools lag behind poor students in other developed countries where the achievement gap between rich and poor is much less prominent (Darling-Hammond, 2011).

As multiple scholars (Baker & LeTendre, 2005; Darling-Hammond, 2010; Turgut, 2013) argue, the U.S. should try to learn from higher achieving nations in order to improve its educational system. This notion of learning from others is one of the underlying ideas that have motivated me to pursue this line of research. Many U.S. educators and policymakers have focused on educational standards, testing and accountability, while largely ignoring the effective implementation of classroom practices from abroad (Kliebard, 2004). The incentive to look outside exists both at the levels of competition as well as inspiration. It might serve U.S. policy makers and educators well to attempt to learn from the successes of other countries—for inspiration and as a way to improve U.S. comparative international rankings on international tests of achievement.

Critics cite issues related to test methodologies, cultural differences, systemic differences and differing rates of childhood poverty as barriers to meaningful interpretation of cross-national comparative tests (Crossley, 2009; Cavanagh, 2012). While most of the nations deemed as high-

performing by test data are more economically and culturally homogenous (e.g., Finland, Singapore, Korea) with very different cultural values, one high-performing nation, Canada, has enough similarities to the U.S. that cross-national comparisons may not have as many inherent difficulties and can be particularly informative (Levin, 2011; Mehta & Schwartz, 2011). The U.S. and Canada have many over-arching similarities including occupational structures, a high standard of living, diverse demographics, decentralization, and geography (Merry, 2013). Furthermore, Gaffield (1994) notes that U.S. and Canadian schools have much in common including a similar history rooted in mass schooling, having developed their schools with similar sociocultural purposes, having populations that include significant socioeconomic, ethnic, and social diversity, and having decentralized local schools without a national curricula. Given these similarities along with the different educational outcomes noted by international tests, I argue that comparing aspects of the U.S. and Canadian education systems [1] is likely to demonstrate differences and [2] that they will derive largely from differences at the school level rather than being primarily attributed to cultural or systemic differences.

Educational Leadership: Towards A Transformation

Most comparative education studies compare schools and nations at the level of policy, curriculum, and demographics (Steiner-Khamsi, 2004; Phillips, 2006; Schreiwer, 1992). Furthermore, much U.S. educational reform has historically focused on alterations in curriculum (Kliebard, 2004). School and teacher leadership, however, have been shown to be among the most important characteristics of school culture and educational success and yet few studies have

looked at these aspects when assessing differences in education cross-nationally (Leithwood & Jantzi, 2000; York-Barr & Duke, 2004).

Traditional notions of leadership asserted that leaders should be singular agents of change who re-define systems on their own. However, in education and other fields, this top-down approach to leadership has proven to be ineffective for learning organizations. Instead, in order to produce maximal results, transformational leadership in the educational sphere where "leaders and followers raise one another to higher levels" (Burns, 1978) has been shown to produce greater levels of motivation and achievement. Kouzes and Posner (2002) have defined five practices of exemplary transformational leadership which lead to exemplary outcomes: "challenging the process", "inspiring a shared vision", "enabling others to act", "modeling the way", and "encouraging the heart". For the purpose of this study, I will define these five core leadership behaviors as measured by a validated self-reported inventory as "leadership behaviors" (Kouzes & Posner, 2007).

While leadership starts with administrators, transformational leadership means that administrators also empower teachers and leadership becomes more widely distributed in schools. There is some evidence that administrative leadership influences teacher leadership and teacher effectiveness (Hipp, 1996; Leithwood, et al., 2006), but these aspects have not been widely addressed cross-nationally and have not been assessed between schools in the U.S. and Canada. Given the importance of leadership on school culture and educational outcomes (Leithwood & Jantzi, 2005; Chin, 2007; O'Donnell & White, 2008; Robinson, Lloyd, & Rowe, 2008), determining whether and how leadership practices differ in the U.S. and Canada may provide important insights into key school-level differences between these countries, differences

that are both relatively amenable to policy change and that have been shown to contribute to academic achievement.

Problem Statement

Despite the importance of school leadership practices, little is known about cross-national educational leadership differences between the U.S. and nations who perform well on international comparative tests. The aim of my dissertation is to contribute to the fields of educational leadership and comparative education by comparing teacher and administrator leadership behavior across urban U.S. and Canadian districts. This work, by identifying and exploring differences and similarities in teacher and administrator leadership behavior, should enable me to determine the degree to which different aspects of school leadership exist cross-nationally and give important insights into how these leadership behaviors impact schools.

Significance Of The Problem

The United States lags behind other nations including Canada in terms of academic achievement. Canadian students have scored higher than U.S. students in reading, math, and science since the first PISA assessment in 2000 (OECD, 2010). U.S. scores in reading and math are 0.3 to 0.4 standard deviations lower in the United States than in Canada, a difference of about one full year of schooling (Willms, 2004).

School leadership behaviors impact academic achievement and represent an important aspect of schooling which has not been explored between the U.S. and Canada, the nation to whose education system the U.S. system can be most easily compared. To address this issue,

this study is designed to assess whether there are differences between leadership behaviors amongst and between teachers and administrators in large urban school districts in the U.S and Canada. This is important because these leadership behaviors have been shown to positively correlate with a number of academic outcomes and to date there has not been a study that addresses these differences cross-nationally.

Purpose Of Study

The purpose of this study is to determine whether there are statistically significant differences between administrators' and teachers' leadership behaviors in large urban school districts in the U.S. and Canada.

Research Questions

1. Are there differences between administrators' and teachers' leadership behaviors in urban school districts in the U.S. and Canada?

In each of the sub-questions I examine the differences between the self-reported leadership behaviors (hereon referred to as "leadership practices") as measured by the LPI instrument at secondary schools in two large urban school districts, one in the U.S. and one in Canada. I refer to these as "cross district" or "across districts" comparisons. I also perform a second set of comparisons across eight districts, four in the U.S. and four in Canada which I refer to as "between the United States and Canada" or "cross-national" comparisons. It is important to note that the term "cross-national" does not imply that study results make any claims of differences between the U.S. and Canadian educational systems as a whole, but rather difference between the four U.S. districts and 4 Canadian districts that were sampled. The rationale for this

important distinction is that neither the USA nor Canada are nationalized educational "systems"—that is, in the US most educational policy is set at the state (US) and provincial (Canada) level.

Sub Questions

For each of these questions, I seek to address differences both (1) across-districts and (2) cross-nationally

a. What *differences* in leadership practices exist between administrators (principals and vice principals)?
b. What *differences* in leadership practices exist between teachers?
c. What *relationships* exist between administrative leadership practices and teacher leadership practices?
d. What *relationships* exist between leadership practices and teacher empowerment?

The statistical hypotheses that I use to test each of these questions is outlined in the research design section.

Research Framework

As my study merges concepts from the broad fields of international comparative education and educational leadership, my overall topic is vast. To narrow the sphere of the study and focus my studies, I use transformational leadership theory (Burns, 1978; Bass, 1985) as a theoretical framework with which to assess differences between teachers and administrators at urban schools in the United States and Canada. Transformational leadership is leadership where

"leaders and followers raise one another to higher levels of motivation and morality" (Burns, 1978). Bass (1985) developed and presented a formal theory of transformational leadership in which leaders raise their colleagues, subordinates, followers, clients, and constituencies to a greater awareness of important issues. Bass (1985) went on to note that this type of leadership requires vision, self-confidence, and inner strength to stand up for what the leaders sees as right, not for what is popular or is acceptable. Avolio, Waldman, & Yammarino (1991) identified four key behaviors of transformational leaders: idealized influence, inspiration motivation, intellectual stimulation, and individualized consideration.

Transformational leadership is seen to be sensitive to organizational learning, building shared vision, distributing leadership, and building a school culture necessary for current reforming efforts in the school (Barnett, McCormick, & Conners, 2001). This type of leadership can occur at multiple levels (teachers, administrators) and has the potential to transform schools and change their culture and their educational outcomes. This type of leadership is positively associated with principals' effectiveness in implementing a reform agenda (Coad & Berry, 1998; Sivanathan & Fekken, 2002). Furthermore, transformational leadership has the potential for building a high level of commitment in teachers in relation to the complex and uncertain nature of the school reform agenda as well as fostering the capacities teachers need to respond positively to this agenda. Leithwood (2007) has demonstrated that transformational leadership stimulates teacher collaboration and commitment thus inspiring individuals in an organization to work together.

The work of Jim Kouzes and Barry Posner (1999, 2002, 2005), who created the LPI instrument as a measure of leadership practices based on their research on effective leadership

over the past twenty years, is based on transformational leadership theory. Their model suggests that leadership is a collection of behaviors and practices and the includes 1) modeling the way, 2) inspiring a shared vision, 3) challenging the process, 4) enabling others to act, and 5) encouraging the heart (Kouzes & Posner, 2005). These five practices are the core domains of the self-reported LPI instrument that I administered. By using questions that measure these specific qualities, my research enables me to gain insights into the types of leadership strategies teachers and administrators utilize. Given the importance of leadership on educational outcomes (discussed further in the review of literature), this framework should provide important insights into how cross-school, cross-district and even cross-national differences in leadership at different levels (teacher/administrator) may impact education in the U.S. and Canada. I undertake a quantitative research study because I want to determine differences in leadership behaviors and teacher efficacy amongst large numbers of educators. In order to obtain a maximal amount of information from a large number of individuals in a rapid and anonymous way, I use surveys (from the Kouzes and Posner LPI) as the primary research instruments (Murray, 2003; Dillman & Schaefer, 1998). In order to maximize the reliability and generalizability of my study, I use only previously validated research surveys and questions in relevant contexts (Schram, 2006). I examine individuals' current behaviors and beliefs as opposed to performing an intervention and assessing its effect, therefore I perform non-experimental research that utilizes a combination of descriptive and correlational statistical analyses (McMillan, 2004). Given this design, I hope to first describe and then use the patterns that I observe in a sample of educators to make inferences about the populations from which they were drawn. I performed rigorous statistical hypothesis testing to objectively analyze the data that I collected.

While I acknowledge that there are inherent limitations to quantitative research which can be problematic in terms of interpretation and oversimplification, I believe that the methods I use enable me to precisely understand differences that may exist in the populations I study and that with statistical inference I am able to draw meaningful conclusions about my research questions. While additional qualitative follow-up of my surveys would be informative and enable further explanation of why differences may exist, I obtained enough data to describe different leadership practices in two large urban U.S. and Canadian school districts. This allowed me to generate important new knowledge and advance the fields of comparative international education and educational leadership.

Vantage Point

From experience as a student and a teacher in the United States and abroad, I perceive that U.S. educators and education policymakers would likely benefit from learning from other countries. In this dissertation, I will use statistical hypothesis testing to determine the validity of null hypotheses, defined as "a prediction that there is no difference between the populations that the two samples are designed to represent", and then determine if any important differences do exist and characterize them using the data I collect (Gall, Gall, & Borg, 2010, 197). This statistical approach is by nature agnostic and prevents me from making statements about differences or conclusions about schools or countries unless there is data to suggest the conclusions that I draw.

For my review of literature, it is my personal and professional educational experience that leads me to ascertain that in order to improve anything at all, including education in the U.S., one

must not focus on learning solely from one's mistakes or the mistakes of people in similar environments or circumstances. Instead, I have found, it is of great benefit to explore and learn from new strategies, different people and various environments and specialized fields. This is precisely why in addition to focusing on supporting literature from scholars whose focus is on education the United States, I also include scholars who extend their comparisons and analyses abroad. To strengthen my diverse and international perspective on learning from other fields, the literature in my research also integrates supportive materials from the arts, business, philosophy, and medicine. I use educational examples from scholars and scientists who have used multiple research designs within qualitative, quantitative, and mixed methodologies. Furthermore, I review a variety of leadership strategies and explain how different types of school leadership may impact educational outcomes.

Organization

My literature review begins with an explanation of what is meant by international competitiveness, and assesses how educational preparedness affects other outcomes in the inter-connected world. Barriers are presented next in the form of curricular struggles, achievement gaps, and barriers of innovation due to standards and accountability. Next I focus on two major international assessments (TIMSS, PISA) and describe their methodologies and results. In addition, I present readers with lessons that can be learned from international tests and point out their limitations. I conclude by presenting overarching international lessons from abroad in terms of preparing students for the world as well as present specific examples from literature about education in Japan, China, Finland, and Singapore. After explaining some of the inherent

difficulties of international comparisons, I review literature about Canada and why this may represent a nation whose educational system can reasonably be compared with that of the U.S. I then go on to discuss leadership and particularly educational leadership. I explore definitions and types of leadership and focus on transformational leadership and five exemplary practices of transformational leaders. Further, I discuss leadership of administrators and teachers, and their links to each other and to educational outcomes. I discuss how leadership impacts teacher empowerment, satisfaction, and self-efficacy. Finally, I address how comparing educational leadership in the U.S. and Canada is both a reasonable pursuit and a worthwhile undertaking. I conclude my literature review by summarizing and contextualizing the body of work that I have reviewed to make clear the place that the study I undertook will occupy amongst the existing scholarship.

B. REVIEW OF LITERATURE

International Tests And Competitiveness

History of U.S. Educational Competitiveness

As many historians note, U.S. political leaders became increasingly concerned about international competitiveness when the Russians launched Sputnik in 1957. This launch began a broader race that also led U.S. political leaders to worry about US schools' international competitiveness and ability to prepare students (Baker, 2007). U.S. leaders interpreted the Soviet Sputnik launch as meaning that scientists from abroad were outperforming their U.S. counterparts, a situation concerning to U.S. policymakers interested in maintaining hegemony. The U.S. began to participate in national tests of academic achievement in 1964 with the First International Mathematics Study (FIMS). Disappointing results on this assessment, along with the release of the Coleman Report in 1966 which demonstrated the importance of race and socioeconomic class in academic achievement, led to a new view of American schooling. Concern for U.S. students' achievement was addressed specifically in the landmark 1983 *A Nation at Risk* report. Specifically, the American education system was heavily criticized for its uninspiring performance in preparing students academically to compete internationally (The National Commission on Excellence in Education, 1983).

Improving U.S. schools has become more important as U.S. students must increasingly compete in a globalized workforce for knowledge-based jobs. President Barack Obama recently declared that "it is an undeniable fact that countries who out-educate us today are going to outcompete us tomorrow" (Cavanagh, 2012, 8). While international tests like the PISA and

TIMSS have limitations and should not be viewed as the sole measures of a country's educational success, they are nonetheless important benchmarks and ways for countries to understand their strengths and weaknesses.

Currently there are two major international educational assessments that are administered to students from a large number of nations. The data from these assessments are used to benchmark national educational achievement. These tests, the Program for International Student Assessment (P.I.S.A.) and the Trends in International Mathematics and Science Study (T.I.M.S.S.), assess student success in distinct ways with TIMSS focusing on recall and factual knowledge and PISA focusing on real world application. For PISA, in addition to providing a benchmark for student achievement internationally, additional goals of the program include providing a means for countries to gauge the pace of their educational progress, demonstrating educational achievements, and establishing reform-worthy policy targets (OECD, 2011). So how useful are these tests at fulfilling these goals? According to Cavanagh and Manzo (2009) "the value of international testing depends largely on policymakers' willingness to probe beneath the raw scores to see what the data say about teaching, the performance of subgroups of students, and other factors" (p.14), meaning that while the test scores have some intrinsic value, deeper and more specific analyses allow for more meaningful interpretation of the results.

Much as happened with earlier comparisons, according to Turgut (2013), the United States' relatively average/modest rankings in current international tests such as TIMSS and PISA are used by politicians as the driving force and rationale for many current educational reforms in the United States in an attempt to "race to the top" of the international rankings. Similarly, Darling-Hammond (2011) warns that a failure of U.S. schools to implement reforms to help

improve education for every child will lead the U.S. to "devolve into a third class power because we have neglected our most important resource for creating a first-class system of education" (p.53).

Results Of International Tests

TIMSS Methodology

The Trends in International Mathematics and Science Study (TIMSS) is administered by the International Association for the Evaluation of Education and measures math and science skills in students from grades 4 and 8 from roughly 50 countries. According to the writers of the TIMSS test, its goals are to monitor system-level achievement trends in a global context, establish achievement goals and standards for educational improvement, to stimulate curriculum reform, and to improve teaching and learning through research and analysis of the data (TIMSS and PIRLS International Study Center,2013). It uses short, fact-oriented stems and primarily multiple-choice questions. TIMSS testing is structured around curriculum based/proficiency benchmarks and specifically focuses on the three cognitive domains of knowing, applying, and reasoning to determine how students have mastered the factual and procedural knowledge taught in school mathematics and science curricula (Kell & Kell, 2010). The test has been administered in 1995, 1999, 2003, 2007, and 2011. While sampling is not uniform worldwide, in the United States, a national probability sample is drawn for each test, this included 480 schools and almost 19,000 students in 2003 (Kell & Kell, 2010).

PISA Methodology

PISA, the Programme for International Student Assessment, is a test designed by the Organization for Economic Cooperation and Development (OECD) that tests 15 year-olds from across the globe in the areas of reading, mathematics, and science. Every three years since 2000, over 65 countries that have participated in these assessments which use open-ended questions that look at real-life applications (OECD, 2007). The test specifically attempts to determine how students can extrapolate from what they have learned with higher order reasoning and analysis to apply their knowledge to novel situations (Schleicher, 2009; Kell & Kell, 2010). Around 470,000 representative students were selected to participate in the 2009 test, representing about 18% of the 26 million 15-year-olds in the schools of the 65 participating countries (OECD, 2011).

TIMSS Results In Brief

The 1995 TIMSS test revealed that U.S. fourth-graders performed only slightly above the international average in both mathematics and science, and the test has shown that U.S. education is not elite in these areas (Baker & LeTendre, 2005). U.S. eighth-graders performed near the international average in both mathematics and science, while the U.S. twelfth-graders scored below the international average and among the lowest of the TIMSS nations in mathematics and science general knowledge, as well as in physics and advanced mathematics (Gonzales et al., 2008). On repeat testing in 2011, U.S. 4th and 8th grade students performed slightly better than previously and above average in both mathematics and science (National Center for Education Statistics, 2012). East Asian nations were the countries scoring highest on

the 2011 TIMSS test including Singapore, Korea, Hong Kong, Taipei, and Japan (TIMSS and PIRLS International Study Center, 2012).

PISA Results In Brief

On the 2006 PISA science assessment, the United States ranked 21st among 30 OECD countries, with a confidence interval that extends from the 18th to the 25th rank (OECD, 2007). Moreover, while the proportion of top performers in the United States was similar to the OECD average, the United States had a comparatively large proportion of poor performers: 24.4% of United States 15-year-olds did not reach Level 2, the baseline level of achievement on the PISA scale at which students begin to demonstrate the science competencies that will enable them to participate actively in life situations related to science and technology (Schleicher, 2009). Given this high rate of low achievement and the large population of the U.S. compared to other nations, the United States produces the most low-achieving students of any of the 34 nations tested (Petrilli & Scull, 2011). PISA's 2009 test showed that U.S. 15 year olds performed around the average in reading (rank 14), and science (rank 17) and below the average in mathematics (rank 25) among 34 participating countries while Korea and Finland were among the top 3 scoring nations in each of these content areas (OECD, 2011). Similarly, the 2012 PISA test demonstrates that students from the United States again performed below OECD average in mathematics, and near the OECD average in reading and science while nations such as China (Shanghai), Japan, Finland, Singapore, and Canada each were each above average (in the top twelve out of sixty-five participating countries) in all three testing categories (OECD, 2013). One caveat about interpretation of US results is that they were not uniform; while some states like Massachusetts

performed near the top of OECD countries, other states (like Mississippi and California) performed very poorly.

U.S. Public Views Of The Educational System

A report summarizing 20 years of American students' performance on the National Assessment of Education Progress (NAEP) tests concluded that "the result is a bleak portrait of the status of student achievement in the United States. Large proportions, perhaps more than half, of our elementary, middle school, and high school students are unable to demonstrate competence in challenging subject matter in English, mathematics, science, history, and geography" and "fewer than 10 percent appear to have both an understanding of the specialized material and ideas comprising that curriculum area and the ability to work with these data to interpret, integrate, infer, draw generalizations, and conclusions" (Mullis, 1990, 29). While Ravitch (2013) notes that U.S. students have made strides in math and reading on NAEP in the last two decades, other scholars still see significant need for improvement. For example, Payne (2008) points out that many of America's urban schools are entrenched in a culture of failure while Darling-Hammond (2010) points out that US schools are unable to provide enough skilled workers for the changing and increasingly information driven economy. In light of these and other issues, a report by Mullis (1990) concluded "the current levels of student achievement are unacceptably low for our country's need and aspirations and for the personal goals of its citizens" (p. 29).

Despite evidence supporting ongoing educational issues, the U.S. public remains more satisfied than they should on the basis of objective evidence (Kinder Institute, 2013). As the

Kinder Institute conducted their study in Houston, for example, residents (and particularly parents) were far more likely to rate their local urban schools as effective, improving, or well-funded than the schools actually were. Surveys of 1000 U.S. parents showed that 71% rated their child's school an "A" or a "B" and 72% were confident in the nations' teachers (Bushaw & Lopez, 2014). Likewise, a Joyce Foundation study found that 76 percent of parents rated the quality of education at their child's current school as good or excellent, that 60 percent believe that the education their child is receiving is much better or somewhat better than the education they received as a child, and that 57 percent of parents say that local public schools are doing a good or excellent job preparing students for college (Tompson, Benz, & Agiesta, 2013). Interestingly, these surveys noted, public opinion of the nations' public schools as whole was substantially lower than that of local schools with only 18% of respondents saying that U.S. public schools as a whole achieved and "A" or a "B" rating (Bushaw & Lopez, 2014).

Criticisms of International Tests

While there is little debate that nations are becoming more inter-connected economically, educational scholars frequently disagree about whether students from countries with different cultural and economic circumstances can or should be compared in terms of academic success due to inherent differences (Schleicher, 2009). Given this fundamental concern, the entire concept of international testing has been called into question and there have been indictments of PISA and TIMSS on the basis of issues of misinterpretation of data, narrow focus, and issues of generalizability (Bracey, 2009; Boe & Shin, 2005; Kell & Kell, 2010; Berliner & Biddle, 1996; Ravitch, 2013). Considering that the results of the tests indicate that U.S. students have lower

scores on international tests compared to high-achieving nations, scholars who do not see the U.S. educational system as lagging behind other nations have been critical of the data from these tests and argue that they are not useful to U.S. students because they divert attention from more important and pressing issues like lack of resources (Baker, 2007).

International tests can be seen from a variety of lenses; as Boe and Shin (2005) note, it is a matter of interpretation:

> One can pick a particular survey (e.g., TIMSS 1995), subject matter (e.g., mathematics), and grade level (e.g., grade 8) and find "many" industrialized nations that scored significantly higher than the U.S. (e.g., France, Japan, and Switzerland). Yet it is also true that U.S. students perform better than students in many industrialized nations. For example, the U.S. scored significantly higher than many industrialized nations (e.g., France, Germany, and Switzerland) in the 1991 Reading Literacy Study at grade 4. Thus, depending on one's interest or agenda, a particular survey result can be selected to support almost any conclusion about how the U.S. stands in the international achievement horse race. (p. 194).

Furthermore, Boe & Shin argue that much of the perception of disappointing U.S. test results is related to math and that this ignores U.S. performance in other subjects.

Zhao (2007) similarly argues that these tests focus too much on rote recall in math and science at the expense of creativity and critical thinking. He argues that if nations focus too much on these tests scores, such a focus can extinguish students' creativity by leading to a teaching to the test mentality instead of one focused on cultivation of students' learning and passion. Furthermore, other scholars argue that the primary data from PISA and TIMSS are

misinterpreted because the tests only measure discrete knowledge and do not assess the quality of nations' educational systems as a whole (Rotberg, 2006). Another issue Boe and Shin (2005) describe is that the way these scores are reported: the U.S. average performance is often mistakenly called poor because U.S. politicians and educators assume that the U.S. should be first in the world academically - a view in which anything less than first place is regarded as a poor performance.

The focus on testing and scores is another concern. According to Kell and Kell (2010), there is a so-called "PISA effect" in which international testing has led to a focus on identifying and quantifying outcomes and outputs; this policy rationale which they argue is a "reductionist approach" reduces education to "attainment of narrow performance and outcome statements" (p.492). This shift in focus, they maintain, serves to suppress the more important questions about what should be taught, why it should be taught, and how it should be taught (Kell & Kell, 2010).

Another major concern with regards to international tests is the question of whether academic achievement can even be compared cross-nationally. Bracey (2009) is one of many scholars who believe that international comparisons are on shaky ground because international tests like TIMSS and PISA are "blunt instruments" which run the risk of making poor comparisons because of differences in each country's different student body and cultures (p.35). Turgut (2013) asks that "before interpreting test results as absolute truth and starting reform efforts based on these tests, they first should be evaluated not only for their technical but also for their cultural and societal validity and reliability" and goes on to state that in order to be applicable, they must "assess what the United States emphasizes in its educational goals and

culture" (Turgut, 2013, p.70). However, while wealthier nations and wealthier students do sometimes outscore poorer ones, sub-analyses of the test data which were conditioned on students' socioeconomic status show that the effect of higher versus lower income on test scores was modest in the 1990s TIMSS tests and not a factor in the PISA 2003 data (Baker & LeTendre, 2005; Wobman, 2007). Baker & LeTendre (2005) note that part of this effect is based on cultural differences and part comes from families investing time and money into their children's education. While one might argue that poverty and diversity are major confounding factors that are more relevant to the U.S. than other countries and might account for some of the achievement gap, this appears to be true mostly within countries rather than between them as TIMSS data show that "classroom diversity in student background, religious-cultural background, and so forth are not associated with cross-national achievement" (Baker & LeTendre, 2005, p. 172).

Another very concerning aspect is the notion that testing leads to teaching to the test and that in doing so, educators will sacrifice more important aspects of education. For example, while Shanghai consistently produces the top scorers on PISA tests, Zhao (2014) notes that China may have the "best education system" because it can produce the highest test scores, but it also has the "worst education system in the world" because those test scores are purchased by sacrificing important aspects such as creativity, divergent thinking, originality, and individualism.

Despite these concerns, international comparisons should not be disregarded. I think that many of the concerns are likely to be over-stated given that PISA uses open-ended questions without one right answer and because it is given in only a limited number of schools and then

used as a national educational barometer but not as a high-stakes test. While any international comparison will necessarily have to deal with cross-national differences which cannot be entirely accounted for, PISA and TIMSS use rigorous methodologies and do not have external incentives to make one country look better or another worse. While all comparisons can be rightly criticized at some level, these comparative tests certainly convey enough valuable information that the arguments of critics like Kell and Kell (2010), Ravitch (2013), and Zhao (2014) do not, in my opinion, substantially invalidate their results.

Educational Equity

A number of other nations have educational systems that produce better results and more equity in test results across students of different socioeconomic status than does the U.S. system. Students from poorer countries like Estonia and Poland outperform U.S. students on standardized achievement tests (Finn, 1997, OECD 2012). The U.S. is not the only country with a long history of injustice and inequality and can learn from some of the measures that countries like South Africa and Brazil are taking to try and address their problems, although certainly both countries are still struggling with this matter (Kubow & Fossum, 2003, 107-140). However, the U.S. system is definitely unique in representing a wealthy country with high overall educational expenditures per capita (McAdams, 1993), a large achievement gap, high childhood poverty rates, and poor measures of academic achievement both internally and internationally. There are multiple factors that are cited as to why U.S. students do not fare well when compared with their international competitors. While issues of childhood poverty, wealth inequality, culture, school funding, teacher training, and curriculum design are commonly cited as major barriers to the

success of U.S. students, comparisons like TIMSS cannot clearly demonstrate which if any of these may be most important (Baker & LeTendre, 2005).

Finland, Singapore, and South Korea are three countries that have "built strong educational systems" and have begun with "very little and purposefully built highly productive and equitable systems" all in the time span of "only two to three decades" (Darling-Hammond, 2011, 21), whereas transformations in the U.S. have been, a struggle, cyclical and less effective at producing consistently high achievement or equality of educational inputs or outcomes (Kliebard, 2004). The schools in Finland, Singapore, and South Korea have made significant improvements in their educational systems during the last thirty years and some of their common strategies have included "fund(ing) schools adequately and equitably, and add(ing) incentives for teaching in high-need schools" (Darling-Hammond, 2011, pp. 21-23,53). While these countries' socioeconomic issues are different than those of the U.S., their success in the area of educational inequity suggests that providing funding and improving staffing at low-performing schools could have positive effects.

A criticism in the literature regarding educational equity and outcomes deals with the issue of childhood poverty. Ravitch (2013) notes that social scientists attribute at least as much importance to childhood poverty rates as they do to schools and teachers when it comes to determining educational outcomes including test scores. This type criticism has also been used to argue that U.S. students have a high rate of poverty (23% according to UNICEF (2012) compared to 13% in Canada) and that this explains why U.S. students lag behind other nations in terms of educational equity. This argument, while important, only holds true across entire nations. Within poorer or urban districts, there still remains a large gap between U.S. students

and similarly disadvantaged students in other nations (Petrilli & Scull, 2011). For this reason, my research will specifically address differences between large urban public schools in the U.S. and Canada which have similar socioeconomic circumstances.

The Role Of Culture

As Cavanagh (2012) notes, culture is a huge driver in educational success: "any single-policy analysis … fails to take into account how great a role cultural norms play in shaping the effectiveness of educational strategies in high-performing nations" (p. 8). He specifically cites differences between U.S. and Japanese students in terms of their cultural expectations and notes that "when U.S. officials look at teaching methods in Japan, they're often surprised by the extent to which educators in that country allow students to struggle with problems, rather than help them"; because while Japanese culture appreciates struggle whereas American parents studied viewed making children struggle to be "torturing my kid" (p. 9). Cavanagh and Manzo (2009) argue that while cultural differences exist, they should not be used as excuses for test scores and prevent policymakers from learning important lessons from international comparisons.

Crossley (2009) notes that using national educational rankings on comparative tests can be damaging because low-performing nations often try to copy high-performing nations without considering the impact of cross-cultural differences that might preclude the higher performing nations' strategies in the lower performing nation. Simola (2005) further notes that because cross-national tests have embedded pedagogical assumptions and values, the test results can prioritize certain types of learning (such as memorization) which run contrary to the types of learning that are most useful to modern students.

Sobe and Kowalcek (2012) note that in order to understand different educational systems, context must be considered and is inherently "caught up in the mobilization of norms, power relations, regulative principles, technologies, and strategies" (p 6). Given the complexity of context, they argue that in cross-national comparisons, researchers must consider the relationality of contexts as "an assemblage of multiple, at times paradoxical, things and practices that come together in particular places at particular times" meaning that to understand an outcome in any educational system, one must also understand the complex circumstances that explain how it came to be. Steiner-Khamsi (2012) further states that to compare systems requires both intra-educational and intra-cultural comparisons which requires investigation of schooling at multiple levels including the historical, social, cultural, political, religious, economic, and philosophical. Since any of these comparisons is a huge undertaking, attempts to compare schools cross-nationally often fail to address some of these important fundamental differences.

In order to better understand the national context, Bray and Thomas (1995) explain that researchers need to employ a multilevel analysis in order to analyze different facets and layers within a system. They describe a cube with geographic, demographic, and social sides which all need to be analyzed in order to understand cultural contexts. Bray and Adamson (2007) note that in order to do meaningful comparative education, researchers should consider context as ethnographers, and as such, should address the wider context in which cultural context produces the educational institutions that are studied.

The above concerns are especially relevant when trying to compare countries with widely different cultures, populations, and contexts. However, these concerns about how differences in culture and economics and social inequity can confound international comparisons also

demonstrate that a carefully designed international comparison is more valid when comparing nations with similar cultures and similar levels of economic and social inequity.

Leadership

Another key factor which is not well assessed by international testing is school leadership: "conspicuously missing from this line of international assessment literature is the connection between school-level leadership" because "constructing a theoretical framework to capture empirical evidence for the linkage between school leadership and student learning is challenging within any learning context, let alone cross-nationally" (Dong & Cravens, 2011, p. 17). However, despite the difficulty of establishing the nature of the relationship between school leadership and student learning in cross-national comparisons, as I will show in section II, school leadership is well established as one of the most important school-level factors in effecting not only teacher morale, but also student achievement (Leithwood & Jantzi, 2005, Robinson, Lloyd, and Rowe, 2008).

In this section, we have seen that the U.S. has not performed well on international comparative tests. While these tests have limitations related to their administration and interpretation, they remain valuable measuring sticks. The tests focus on math, reading and science and do not measure "flexibility, openness, and tolerance" which are American strengths; however, these tests are the best basis we have for comparing international achievement (Zhao, 2007, p. 14). In addition to achieving more on tests, high performing countries have made strides for equal and fair distribution of resources so that disadvantaged students do not suffer disproportionately to others. Finland, Singapore, and Japan serve as three examples of countries

that have been able to achieve this level of success and balance in education. While the U.S. cannot become these nations, it can use them as exemplars and learn from them. As Darling-Hammond (2010-2011) stresses, "while no system from afar can be transported wholesale into another context, there is much to learn from the experiences of those who have addressed problems we encounter" (p.21). Canada may represent the nation whose schools are most similar to those of the US and therefore it may represent the best nation with which to compare US schools. For this reason, I will next look to Canada as a comparator nation.

Canada

Fair Comparison To The U.S. Educational System

The U.S. and Canada are similar in the sense that they share a continent, occupational structures, a high standard of living based on GDP, and are demographically diverse (Merry, 2013). Furthermore, Gaffield (1994) explored the history of U.S. and Canadian schools and found that the two nations' education systems share many common origins, philosophies, and structures. Levin (2011) argues that Canada of today continues to have many similarities to the U.S. in education including little federal government presence in schooling, local districts with elected boards, and high levels of student diversity, particularly in urban schools. Furthermore, US and Canadian schools have similar curricular requirements, types of schools, grade levels, and graduation rates (Levin & Sagedin, 2011).

Mehta and Schwartz (2011), demonstrate how a country very similar to the U.S. in terms of economic and cultural diversity has succeeded using different educational policies and strategies. Regional curricula, training and selecting the top teachers, a national reform agenda,

and equalizing funding are among the top Canadian goals much like the other high-achieving nations. Compared to the U.S., Levin (2011) notes, the main difference is that Canada has a much smaller proportion of low performing students with performance at the top end of the distribution being similar. This difference likely accounts for at least some of the differences seen between these nations on comparative tests like PISA and TIMSS. By focusing on urban schools where there are large numbers of low-income students, my research design importantly explores some of the differences that may contribute to Canada's smaller number of low performing students.

Levin (2011) points out important differences in education between Canada and the U.S. include that Canada has (1) better trained, more respected, and better paid teachers (2) a commitment to equity for all population groups (3) more comprehensive basic services such as health care and social services (4) smaller differences in funding levels from one district to another (5) generally more spending in higher need communities and (6) greater consistency across schools and districts in curriculum and teaching methods. Interestingly, Levin notes, these differences arise largely from different treatment of education and schooling at attitudinal and local/provincial school policy levels rather than from national policy or cultural differences between the U.S. and Canada. Unlike circumstances where there are significant cultural differences and this nearly precludes cross-national educational comparisons such as Cavanagh (2012) describes between the U.S. and Japan, the differences that Levin (2011) cites between the U.S. and Canada seem to pose much less significant barriers to comparing these nations' educational systems.

Canadian Test Outcomes

Canadian students consistently scored higher in mathematics, reading, and science as compared to U.S. students on the PISA test (OECD, 2013). In addition to the six differences between the U.S. and Canada cited by Levin, three other factors which help explain Canadian educational success are cultural factors like high levels of parental support, Canadian embrace of a welfare state and thus a sense of need to give a high quality education to all students, and policy factors which have led to a pool of highly qualified teaching candidates and funding which is equalized based on student neediness (O.E.C.D., 2010). Indeed, Canada was the sixth highest scoring nation overall, the highest English-speaking and highest scoring French-speaking nation which has made Canada a "go-to" country for educational inspiration and policy learning (Hargreaves & Shirley, 2012). Given the major similarities between U.S. and Canadian schools including their geography government, school structure, and diversity, these authors conclude that the successes of Canadian schools are due to not only policy but also local attitudes including high regard for public education, support of teachers, and collaborative and inclusive processes of educational change management, (Hargreaves & Shirley, 2012).

Sources of Improvements in Canadian Schools

While changes have been local and incremental, some Canadian districts have recently begun to make both attitudinal and structural improvements. Tucker (2011) notes that many Canadian schools have improved significantly largely in part to the fact that after making their financing more equitable, they have depended on their teachers to raise the achievement of

students who were previously doing least well. Furthermore, students who are not engaged by the academic curriculum have a "different menu of courses" and are able to gain valuable skills that lead to employment opportunities (Tucker, 2011, p. 157).

Many of the policy-level changes noted by Tucker have been rooted in Canadian school leadership paradigms. Fensterwald (2013) notes that Canadian schools have overcome "stagnation" by developing a partnership between teachers and administrators to help improve achievement. Improvements implemented in Ontario, she notes, helped to establish a "collaborative culture to get teachers to work together, led by principals who know how to focus on instruction." In this way, Ontario teachers were held to greater levels of accountability but using innovative strategies such as transparency of results coupled with non-judgementalism that allows administrators to help struggling teachers without a sense of competition.

Defining Leadership: Towards the Transformational and Operational

Leadership in the Educational Sphere

Leithwood and Riehl (2005) define educational leadership as ''the work of mobilizing and influencing others to articulate and achieve the school's shared intentions and goals" (14). They further go on to describe successful school leadership as ''leadership orientations and practices that have been demonstrated to have a positive impact on student learning, whether directly or indirectly through school conditions or the actions of others" (14-15). They report multiple practices which help achieve this goal including making important contributions to the improvement of student learning, distributed of leadership across the school and school

community, a core set of leadership practices, and the enactment of practices to promote school quality, equity and social justice through building powerful forms of teaching and learning. Interestingly, there are few cross-national studies of educational leadership, much of the scholarship in this field is produced by U.S. scholars such that comparatively less is known about leadership strategies and practices in other nations (Oplatka & Addi-Raccah, 2009).

Toward An Operational Definition Of Leadership

Understanding leadership as a set of core set of practices makes it much easier to evaluate leadership objectively. Since "leaders do exhibit certain distinct practices when they're doing their best", Kouzes and Posner set out to define specific behaviors which defined exemplary leadership. To this aim, Kouzes and Posner (1995 & 2002) developed a model for leadership based on five practices of exemplary leadership. Each practice contained two commitments woven into a core theme. Kouzes and Posner (1995), after examining thousands of "personal-best" experiences, developed a quantitative instrument called "The *Leadership Practices Inventory*" (LPI) which assesses specific domains of leadership. In 2002, they codified the following five practices of extraordinary leaders which serve as guidance for leaders to accomplish their achievements or "to get extraordinary things done" (Kouzes & Posner, 1995): challenging the process, inspiring a shared vision, enabling others to act, modeling the way, and encouraging the heart (Kouzes & Posner, 1995, 2002). Given the utility of this method, I will define leadership for the rest of this work as the performance of these five exemplary practices. The leadership measured by the LPI has been noted to be primarily transformational leadership.

Five Practices Of Exemplary Leaders: The Kouzes And Posner Leadership Model

To quantitatively assess leadership there needs to be a valid, reliable assessment procedure. Kouzes and Posner (2007) have, through extensive study, found that leadership is not about personality, but about behaviors (Kouzes & Posner, 2007). Their research suggests that leadership is not a position, but a collection of practices and behaviors. These practices serve as guidance for leaders to accomplish their achievements or "to get extraordinary things done" (Kouzes & Posner, 1995, p. 9). They found that successful leaders engage in the five core leadership practices named above and described in detail below (Modeling the Way, Inspiring a Vision, Challenging the Process, Enabling Others to Act, and Encouraging the Heart). After a rigorous process and validation, these practices were turned into questions and make up the Leadership Practices Inventory (LPI), one of the two survey instruments used in this study. The five practices of exceptional leaders are each individually measured by the LPI subscales and examine different aspects of leadership all of which are relevant to the school setting.

(1) Model the Way.

According to Kouzes and Posner (1997), transformational leaders set an example and build commitment through daily acts that create progress. In order to gain trust from their organization, they create a program of excellence in which they participate and set an example for others to follow. In this way, leaders show that they live by the values they advocate (Kouzes & Posner, 1995, 2002). Exemplary leaders model behavior consistent with their beliefs and their expectations of others. Indeed, modeling the way is based on core values: "To effectively model the behavior they expect of others, leaders must first be clear about guiding principles" (Kouzes & Posner, 2007, p. 15). Leaders give voice to their values, set an example through their daily

actions, and earn the respect needed to lead through direct involvement. Jacoby (2004) notes that as key stakeholders in educating youth, school leaders must not only strive for excellence in their personal work but must also pursue that excellence in the character of their leadership.

(2) Inspire a Shared Vision

Kouzes and Posner (2002) note that leaders have visions and dreams of their future ideal organization and that one of their jobs is to share their vision. Leaders inspire the hopes and dreams of others, forge unity of purpose, stir passion, and "communicate their passion through vivid language and an expressive style" (Kouzes & Posner, 2007, p. 18). Inspiring a shared vision brings people in an organization together to commit to a shared future. Furthermore, this type of leadership fosters enthusiasm for a common vision through both passion and by and skillful use language and energy (Kouzes & Posner, 1995, 2002). In schools, when there is positive rapport, trust, and respect between teachers and administrators, the likelihood of improved pedagogy and increased student achievement is substantially increased (Zimmerman & Deckert-Pelton, 2003).

(3) Challenge the Process

By either creating new ideas and recognizing and supporting new ideas, challenging the process is a practice that enables leaders to show a willingness to oppose the status quo in order to turn ideas into actions (Kouzes & Posner, 1995). Moreover, leaders who seek out challenges test their skills and abilities and often find innovative ways to improve their organization(s). This process of challenging current practice leads them to experiment and take risks. While this means that transformational leaders must be prepared to make mistakes, mistakes in this context function as part of a learning process. Instead of punishing failure, leaders who utilize this

practice encourage failure and to learn from mistakes rather than shift the blame on someone else (Kouzes & Posner, 1995, 2002).

(4) Enable Others to Act

Exemplary leaders need to foster collaboration and build trust. Kouzes and Posner (2007) found that leaders who make their constituents feel strong and capable are able to create organizations where people want to give their all and exceed expectations. They further note that authentic leadership is founded on trust, and that when there is trust in a leader, members of an organization are more likely to take risks, make changes, and foster organization and movements. Transformational leaders strive to create an atmosphere of trust and dignity to help each person in the organization feel capable of acting (Kouzes & Posner, 1995).

(5) Encourage the Heart

According to Kouzes & Posner (2007), "It's part of the leader's job to show appreciation for people's contributions and to create a culture of celebrating values and victories" (p. 22). Along similar lines, Sergiovanni (2005) notes that, "strengthening the heartbeat of the organization is key to building a culture of leadership and learning" (p. 2). People often need encouragement and motivation to achieve the goals set by their organization. By influencing motivation, leaders attach rewards and recognition to job performance (Kouzes and Posner 2002). Transformational leaders play a special role in the celebration of individual and group achievements because they serve as a role model. By celebrating achievements together, leaders help people feel that they are part of the group and a part of something significant. Because leaders encourage their employees through recognition and celebration, they have the ability to inspire them to improve performance (Kouzes & Posner, 1995, 2002). Furthermore, when

leaders use celebrations and rituals to build a collective identity, this can help carry a group through tough times (Kouzes & Posner, 2007).

Educational Studies Using The LPI

While there are thousands of studies which have used the LPI in business settings, it is also applicable to schools. Richardson, Flanigan, Lane, and Keaster (1992) surveyed 1,225 teachers with the LPI in an effort to determine teachers' perceptions of desirable characteristics of principals' leadership behavior. The rankings of teacher perceptions of principal characteristics were very similar to the rankings of business employees for their managers. Both groups perceived that leaders were most concerned that their leaders demonstrate honesty. Other characteristics ranked highly by teachers were competency, forward-looking, inspiring, and caring.

Starcher (2006) conducted a study to determine if a significant relationship existed between the leadership practices of school principals and student achievement in mathematics and reading. The study surveyed 350 principals who had served as principals of their schools for three years or more. A significant relationship was found between the leadership practice of "model the way" and cases of student achievement in reading.

More recently, a study by Leech and Fulton (2008) used the LPI to examine relationships between teachers' perceptions of the leadership behaviors of principals in a large urban school district and their perceptions of the level of shared decision making practiced in their schools. Participants in this study consisted of 646 teachers from 26 secondary schools in a large public school system. Each of the participants was given two survey instruments, the LPI and the

Shared Educational Decisions Survey which measure shared decision-making. This study found significant relationships between the leadership behaviors of principals and the level of shared decision-making. The strongest relationship was between the leadership practice of "Challenging the Process" and the level of shared decision-making in the area of policy development (Leech & Fulton, 2008).

Leech, Smith, Green, and Fulton (2003) surveyed 242 teachers in a large urban school district and asked them to assess their principals' leadership practices. Over half of the respondents perceived that their principal demonstrated effective leadership practices "fairly often" to "almost always." Principals most often exhibited the practices of "Enabling Others to Act" and "Modeling the Way", while they rarely demonstrated "Encouraging the Heart" (Leech, Smith, Green, & Fulton, 2003).

Principals And Teachers As Leaders

Principals As School Leaders

According to traditional educational leadership dogma, the most important factor in school effectiveness is the principal (Edmonds, 1979; Hallinger & Heck, 1996) and the instructional leadership provided by that person (Heck & Hallinger, 1999). A principal, as the school leader, is often viewed as instrumental in developing and inspiring teachers as that person sets the tone and direction for a school (Leithwood & Jantzi, 2000).

According to Glickman (2002), the key functions of an administrator (principal) are the supervision and the leadership of the teaching staff. This includes establishing a professional working and learning environment, supporting teachers' professional growth, and providing

resources to teachers (Dipaola & Hoy, 2008; Zepeda, 2007). Principals are tasked with helping teachers develop their skills so that they can better facilitate student learning. Fink and Resnick (2001) argue that principals help teachers make curriculum choices, establish expectations for student work, and provide teachers the opportunities to learn the specifics of teaching effectively within their academic areas.

A growing body of evidence has demonstrated a significant and positive relationship between principal leadership and student achievement. Recent research including qualitative (Madden, 2001) and quantitative studies examining indirect leadership effects on student outcomes (Heck, 2000; Leithwood & Jantzi, 2000; Marks & Printy, 2003) conclude that although teachers are on the frontlines of learning and have the greatest impact on student achievement and motivation, the school principal's leadership has a vital impact on school culture (Ross & Gray, 2006). Robinson, Lloyd, and Rowe (2008) point out that school leaders who engage in activity connected to the classroom are more likely to positively influence student learning outcomes.

Transformational leadership amongst administrators can provide intellectual direction and innovation within the organizations while empowering and supporting teachers as partners in decision-making (Leithwood & Jantzi 1994, 2000). Leithwood and Jantzi (2006) conducted a study on the effects of transformational leadership on teachers' motivation, capacity, working settings, classroom practices, and student achievement. They showed that principal leadership directly influenced teachers' motivation and significantly influenced teachers' capacity. The specific actions that principals engage in to do this include the establishment of a positive school culture (Bason & Frase, 2004), modeling of expectations by school leadership (Brooks et al.

2007), and building teacher leadership through the impact of the teacher's perception of leadership decisions (Noonan & Walker, 2008).

For these reasons, studying principals/administrators' leadership is crucial to fully understand the leadership culture and practices in a school. For my study, one of my goals focused on investigating administrators' leadership behaviors and determining whether they differ between U.S. and Canadian schools.

Teachers as Leaders

While teachers have always held critical school responsibilities, the concept of teacher leadership as being critical to school success arose in the 1980s with scholarship which emphasized operational leadership models and the importance of active involvement of individuals at all levels in making important decisions (York-Barr & Duke, 2004). It is difficult to establish the exact origin of the concept, and there has been substantial debate about what "teacher leadership" means (Barth, 2007). As a consequence, the research on teacher leadership is somewhat fragmented due to the lack of clarity of the concept itself and a lack of empirical research into teacher leadership (Jackson, Burns, Bassett, & Roberts, 2010). Some of the literature focuses on the elements of teacher leadership or the role of teacher leadership in educational improvement. For example, York-Barr and Duke (2004) note that:

> Recognition of teacher leadership stems in part from new understandings about organizational development and leadership that suggest active involvement by individuals at all levels and within all domains of an organization is necessary if change

> is to take hold…. Educational improvement at the level of instruction, for example, necessarily involves leadership by teachers in classrooms and with peers.

A report by the Coalition of Essential Schools noted that teachers who self-identify as teacher leaders rarely make up more than 25% of a faculty (Barth, 2001). Yet, studies of teacher leadership have found that teacher leaders can influence policy at the district level and make a difference at the school level through their expertise (Hatch, White, & Faigenbaum, 2005), and that support of teacher leaders is critical for school reform to occur (Silins & Mulford, 2004). Moreover, Ryan (1999) found that teacher leadership brings about school change, promotes democratic schools, and transforms schools into places of adult, as well as student, learning.

Relationship Between Administrative And Teacher Leadership

As shown in the principals as leaders section above, principals are critical to teacher leadership support and success in a school. They recognize a job well done, provide empowerment in the form of decision-making, and share in the responsibility when initiatives fail. Furthermore, principals frame the context in which the teacher leadership process succeeds or fails (Moller, et al., 2001). Thus, as Acker-Hocevar and Touchton (1999) report, schools with high levels of teacher leadership are led by principals who are most willing to share power and release control (distributive leadership). They found that such principals respect and trust teachers and cultivate school conditions supportive and conducive to the effective and empowering practices. Within a school, empowering others to lead alongside the principal builds collegiality and active participation in school improvement. Schools with these cultures are

referred to as learning organizations, characterized by collaboration, risk taking, and a shared mission (Silins & Mulford, 2004).

Administrators play a direct role in the success of teacher leadership practices through their own leadership behaviors (York-Barr & Duke, 2004). Thus, administrators who exhibit effective leadership styles can or may impact the forming of similar leadership styles amongst their schools' teachers.

Teachers benefit from a distribution of power that gives them a voice in school wide decisions (Ingersoll, 2007). Marks and Printy (2003) performed a mixed-method study of 910 teachers and found that student achievement was at its highest when quality teaching was the result of integrated leadership. These authors note that sharing leadership responsibilities enhanced student achievement and caused less burnout in the principals (Marks & Printy, 2003).

Leech and Fulton (2008) performed a correlational study to determine the relationship between teachers' perceptions of the level of shared decision making practice in their schools and their perceptions of the leadership behaviors of their principals. Amongst 646 participants from 26 schools in a large urban public school system, they found there was no significant relationship between the leadership behaviors of the principal and the level of shared decision making in schools. Conversely, Leithwood, Jantzi, and Fernandez (1993) surveyed 168 teachers from nine secondary schools and found that vision-creating and goal consensus-building strongly impacted teacher commitment to change. Also, they noted that leadership practices that created direction, meaning, and purpose to the teacher's work greatly enhanced teacher commitment.

The work of these scholars shows a link between administrative leadership and teachers' commitment to change and involvement in school-based professional learning communities. Given such findings, it seems important to determine the leadership behaviors of administrators in order to understand how these correlate with and potentially affect teachers' attitudes. Furthermore, because the above studies have shown conflicting results as to whether principal leadership behaviors influence teacher leadership, it will be important not only to determine leadership behaviors in administrators but also to correlate these with leadership behaviors amongst teachers.

Distributed Leadership

As opposed to top-down leadership, many authors note that distributing leadership across many individuals leads to greater organizational coherence. Hulpia, Devos and Roseel (2009) note that three core functions of effective school leadership can be distributed: (a) setting a vision, (b) developing people, and (c) supervising teachers´ performance. Through distributed leadership in schools, teachers especially are able to draw on and develop practices which are informed by evidence and research, as well as providing them with a voice in professional and school-based decision making (Veuglers & O'Hair, 2005).

More than simply acknowledging the importance of multiple leaders, the concept of distributed leadership considers leadership practice, which is a product of the interaction of leaders, followers, and the situation (Gronn, 2002). The notion that leadership is "stretched over" people and place (Spillane & Sherer, 2004, page) makes the distributed perspective unique in a leadership landscape that frequently looks to omnipotent leaders that are single-handedly

responsible for the success of their organizations. Muijs and Harris (2003) argue that the notion of distributed leadership refutes the myth of the heroic leader, and in doing so concedes that the authority to lead is not exclusively located in formal positions, but is dispersed throughout the organization across individuals, structures, roles, and routines.

Effect Of Leadership On Academic Achievement

Leithwood and Jantzi (2005) conducted a meta-analysis of both qualitative and quantitative studies on transformational leadership in schools. The authors developed their own set of transformational leadership behaviors which included setting direction, helping people and redesigning the organization. The authors also reviewed mediating variables in which leaders exercised their influence. Transformational leadership was found to have an impact on teacher commitment and job satisfaction. Transformation leadership was also found to have a significant effect on student achievement. There was little evidence of a relationship between the transformational style of the principal and school culture (Leithwood & Jantzi, 2005).

Chin (2007) attempted to determine the significance of transformational leadership as an important aspect of overall school effectiveness. A meta-analysis of 28 studies was completed to inspect the relationship of transformational leadership behaviors and teacher job satisfaction, school effectiveness as perceived by teachers, and student achievement. This study concluded that transformational leadership behaviors had positive effect sizes in relation to teacher job satisfaction, school effectiveness as perceived by teachers, and student achievement.

Higher teacher perceptions of principal leadership also correlate with higher student achievement according to a quantitative correlational Pennsylvania study of 325 randomly selected middle school educators; a significant relationship was shown between how teachers perceive a principal's promotion of the school learning climate and student learning gains in math and reading (O'Donnell & White, 2008).

A study of student achievement and leadership style in rural schools in New Mexico found that student achievement was related to the principal's ratings on participation, role clarification, supervision, and charismatic leadership (Edington, 1988). This study concluded that school principals must make decisions on what to do about student achievement and school effectiveness.

Using the results of PISA and TIMSS as a barometer, the Canadian education system has demonstrated higher levels of student achievement than the U.S. While there are differences in the two countries' politics and educational structure, differences in leadership behaviors' effects on educational outcomes and cross-national differences has not been previously explored. Given the mounting body of evidence suggesting the importance of school leadership in influencing achievement, this study helps determine whether leadership differences at the level of teachers and administrators in these two countries might be a major contributing factor to the gap seen between these two nations' test scores.

Comparing U.S. And Canada's Leadership Is Reasonable

Given the many concerns about cross-national comparisons, it is difficult to draw conclusions from many of the cross-national studies reviewed as to whether these findings may

be relevant to U.S. education, particularly when it comes to leadership. Many cross-cultural studies suggest that both national/societal and local school culture can influence leadership concepts, styles and practices (Gerstner & Day, 1994; Hofstede, 2001; House & Aditya, 1997). As House et al. (1999) suggest, "what is expected of leadership, what leaders may or not may do, and the status and influence bestowed upon them vary considerably as a result of the cultural forces in the countries or regions in which the leaders function" (4). In order to determine whether a country can be fairly compared, Hofstede (1980) assessed nations' individualism which correlates strongly with leadership styles and found that the U.S. and Canada were both "low power distance societies in which people value equality, rights, independent thinking and democratic decision-making (Hofstede, 1980; 2001, page). The Power Distance Index of the United States (PDI = 40) and Canada (PDI = 39) had similar scores suggesting that they reject hierarchical forms of leadership, and have increased acceptance of leadership models such as distributed, shared and transformative leadership compared to other nations.

Canadians Have Implemented Changes In Leadership Practices

Fullan (2008) has initiated systems change through the use of strategies that foster leadership at all levels of the education system in Canada. Through research and direct application in Ontario, he has demonstrated that the supportive and supervisory actions of each positive teacher leader in these schools help to cultivate other teacher leaders, who then begin to collaborate together for a common good (Fullan 2001, 2007, 2008). He further argues that in Ontario, steady increases in performance are partly a result of networking strong schools with weaker performing peers across districts. The provincial government has placed a focus on

providing significant career development for existing school leaders, while also developing distributed capacity and lines of succession within each schools (Hargreaves & Shirley, 2009).

Levin (2008) explains that Ontario principals all have teaching experience and receive mentoring funded by the provincial government and organized through local school boards which include training for mentors as well as defined mentoring plans. Canadian principals then "set a number of challenging goals and strategies to achieve those goals" which "becomes the basis of their evaluation". To ensure leadership is stable, the author continues explaining that principals in Ontario have "a leadership succession plan" to ensure schools' academic gains are not lost when principals leave.

Measures such as these have been effective even in low-achieving urban schools in Canada. In Toronto, Leithwood and Strauss (2009) note that administration focused on literacy and held professional development workshops on this subject which they began to hold teachers accountable for. Rather than resenting this accountability, teachers demonstrated a willingness to work collaboratively with their colleagues, and a sense of school-wide responsibility for student success began to pervade the professional cultures of schools labeled as "turnaround" because a significant improvement in school achievement was required.

Teacher Empowerment And Satisfaction

Empowerment Defined

Empowerment is defined as the opportunities an individual has for autonomy, choice, responsibility, and participation in decision-making in organizations (Lightfoot, 1986). Jenkins (1988) stated that leaders who empower others give stakeholders a share in the collective

movement and the direction of the organization. In the school context, empowerment, according to Short, Greer and Melvin (1994) is defined as, ‘‘a process whereby school participants develop the competence to take charge of their own growth and resolve their own problems’’ (p.38). It is individuals’ belief that they have the skills and knowledge to improve a situation in which they operate.

In their national study on empowerment of teacher leaders, Rinehart and Short (1991) found that teacher leaders were more highly empowered than classroom teachers. This finding was explained as a result of teacher leaders’ having more opportunities to make decisions and grow professionally, having control over their daily schedules and feelings of a high level of teaching competency. According to Maeroff (1988), teacher empowerment consists of improved status, increased knowledge and access to decision-making.

Six Dimensions Of Teacher Empowerment: The SPES Model

Short and Rinehart (1992) used factor analysis of the School Participant Empowerment Scale (SPES) instrument to identify six dimensions of teacher empowerment. In a follow-up study, Short (1994) further describes the six dimensions detailed below: decision-making, professional growth, status, self-efficacy, autonomy, and impact.

1. *Decision-making.* This term refers to teachers’ participation in decisions that directly affect their work. In addition to classroom decision-making, such decisions include issues related to budgets, teacher selection, scheduling, and curriculum. To be effective, teachers’ participation in decision-making must be genuine, and teachers need to be

confident that their decisions actually impact real outcomes (Short 1994). Furthermore, allowing teachers to be involved in decision-making results in teachers being responsible for solving problems, not just identifying them (Short, 1994, Whitaker & Moses, 1990).

2. *Professional Growth.* This category refers to teachers' perception that their school provides them opportunities to grow and develop professionally, continue to learn, and expand their skills, and work collaboratively (Short and Rinehart 1992). A survey of over 1,200 teacher-leaders showed that teachers who have the opportunity to collaborate with their peers improve their teaching effectiveness (Berry et al, 2010).
3. *Status.* This term refers to the professional respect and admiration that the teachers perceive that they earn from colleagues (Short 1994). Respect is also given to teachers for their knowledge and expertise and the resulting support of their actions from others (Short, 1994; Whitaker & Moses, 1990).
4. *Self-efficacy.* This concept is defined by teachers' perception that they are equipped with the skills and ability to help students learn, and are competent to develop curricula for students. The feeling of mastery, in both knowledge and practice, that results in accomplishing desired outcomes is critical in the teachers' sense of self-efficacy. The concept of self-efficacy is drawn from Bandura's social cognitive theory (1977, 1986). Bandura (1986) noted that people's belief in their capabilities to "exercise control over events that affect their lives" (p. 1176) is central to human nature. Bandura (1977) thus introduced the idea of self-efficacy as "beliefs in one's capacity to organize and execute the courses of action required to produce given attainments" (p. 3).

5. *Autonomy.* This subscale measures teachers' feeling about whether they have control over key aspects of their work life. This type of control enables teachers to feel free to make decisions related to their educational environment such as scheduling, planning instruction, and curriculum development (Rinehart 1994). Not only does autonomy make teachers feel more independent, but according to Berry et al. (2010), teachers afforded autonomy in decision-making also become more effective in the classroom.

6. *Impact.* The term impact in this context refers to the teachers' perception that they can effect change and influence their school (Short and Rinehart 1992). To help ensure that teachers feel impactful, schools should provide resources to help teachers feel that they are of value to the school, that they are positively effecting the teaching and learning process, and that their ideas are put into practice (Short 1994, Whitaker & Moses, 1990, and Davidson & Dell, (2003).

Wall and Rinehart (1998) describe that the dimensions of empowerment that are perceived most frequently amongst high school teachers are status, self-efficacy, and impact with lower scores on professional growth, autonomy and decision-making (categories ranked in descending order). This implies that teachers feel that they are respected are effective at their job and are able to accomplish things at school but frequently do not feel that they are able to function independently or that they are involved in the process of school decision-making.

Work by Sweetland & Hoy (2000) makes four assertions regarding teacher empowerment. These include the notion that teacher empowerment is most effective when it is works to increase professionalism, that next, empowerment exists at both the organizational and

the classroom level, that empowering teachers should focus on teaching and learning in schools, and to be authentic (pp. 710–711). Teacher empowerment is, therefore, perceived as a crucial factor that affects school effectiveness (Wall & Rinehart, 1998).

Quaglia, et al. (1991) describe that school administrators can use teacher empowerment and teacher satisfaction to provide opportunities for teachers' participation and leadership within the school. They note that administrators should be trained in the types of leadership that encourage empowerment . Specifically, administrators can focus on creating an empowering work environment for teachers by "increasing participation in decision-making, providing opportunities for professional growth, fostering professional respect between colleagues, promoting self-efficacy, allowing for teacher autonomy, and welcoming teacher input on school related issues" (Quaglia et al, 1991).

Teacher Satisfaction

In addition to the six aspects of empowerment, another factor which needs to be addressed to fully understand teachers' efficacy is satisfaction. Job satisfaction is a multifaceted measure which includes s an overall feeling about one's job or career which relates to specific details including compensation, autonomy, coworkers, and relationship with administration. Perrachione, Rosser, and Peterson (2008) found that multiple variables were significantly related to teacher satisfaction and retention and found that when teachers view the school's policy environment as unfavorable, they are likely to leave teaching.

Teacher job satisfaction also has important implications for student learning. Choy et al. (1993) found that teachers who are not satisfied may be less motivated to do their work while

those who are highly satisfied are less likely to change schools or to leave the teaching profession. Bogler (2001) looked at how the behavior of principals affects teacher job satisfaction. Factors they found to be related to teacher job satisfaction included principals' leadership style, decision-making strategy, and teachers' perceptions of their occupation (Bogler, 2001). The study found that when teachers perceived teaching as a profession as opposed to a job, they were more likely to see their principal as a transformational leader (Bogler, 2001). Moreover, teacher empowerment is positively correlated with job satisfaction (Rinehart & Short, 1994). Conversely, teachers with low job satisfaction perceived their principals as transactional leaders with an autocratic decision making style (Bogler, 2001). Additionally, Whaley (1994) reported results of a study that showed a strong relationship between principal communication and job satisfaction, especially in the areas of feedback, rewards, and support. Given the clear inter-play between administrative leadership and teacher self-efficacy, I utilized focused questions from Short and Rinehart's (1992) validated School Participant Empowerment Scale (SPES) instrument to determine if there are differences in teacher empowerment and efficacy across U.S. and Canadian teachers and how these factors were impacted by administrative leadership behaviors.

Relationship Between Empowerment And Leadership

An important derivative of leadership is the ability to inspire. Specifically, an effective school administrator can impact student learning gains by making decisions that encourage school capacity through a collaborative school environment and by inspiring teachers. At low-performing schools, a major challenge for teachers is low teacher efficacy. According to Nir and

Kranot (2006), studies have shown that teachers' self-efficacy, which reflects a perceived ability to produce a positive improvement among pupils, is one of the most influential factors on the quality of teaching, on teachers' efforts and motivation, satisfaction, and ultimately, on students' academic outcomes. Teachers with high rates of self-efficacy are better able to cope with stress, have a higher commitment to teaching, and are more willing to incorporate new teaching methods. Moreover, these efforts lead students to have higher levels of motivation and diligence (Nir and Kranot 2006). It is therefore important to study administrators' leadership behaviors and determine whether and how they may change teacher self-efficacy.

Teachers' confidence in their ability to perform the duties that motivate students' learning is one of the most important individual characteristics that predicts teacher practice and student outcomes. Teachers need a high level of efficacy beliefs in order to continue teaching at a highly effective level over the life of their careers. Swackhamer, Koellner, Basile, and Kimbrough (2009) indicate that teachers' sense of self-efficacy is an important attribute of effective teaching and is positively correlated with student outcomes including scores on the mathematics section of the Iowa Test of Basic Skills. Furthermore, studies have shown that teachers with high levels of self-efficacy work longer with students that struggle, and are more likely to recognize student errors and attempt new teaching methods that support students (Blase & Blase 1999; Barnett, Craven & Marsh, 2005).

Teacher self-efficacy beliefs can be proliferated through transformational leadership practices (Hipp, 1995). Several leadership practices have proven to influence teacher self-efficacy beliefs. Specifically, principal leadership practices that contribute to teacher self-efficacy include emphasizing accomplishments, increasing teachers' confidence, involving

teachers in school decision making, being responsive to teachers' concerns, promoting an academic emphasis in school, and providing supervision which teachers deem to be useful (Marzano, 2005; Hoy & Woolfolk, 1998).

Hipp's (1996) study of the influence of principal leadership found that principals who adopt transformational leadership practices were more likely to have teachers at their school with high self-efficacy. The study found that there were numerous behaviors that a principal could do to impact teacher self-efficacy and that leadership techniques. These included modeling appropriate behavior, believing in teachers, inspiring group purpose, promoting shared decision-making, recognizing teacher efforts and accomplishment, fostering teamwork and collaboration, and encouraging innovation (Hipp, 1996).

Goddard and Hoy-Woolfolk (2004) found that teacher efficacy provided student learning gains when they studied teachers in 47 elementary schools. Their data showed a positive correlation between student achievement and schools with a positive collective teacher efficacy (Goddard & Hoy-Woolfolk, 2004). Furthermore, achievement rose when teachers believed they were part of a competent staff with the ability to overcome educational obstacles. Students with high-efficacy teachers showed more academic advancement in schools while teachers with a high sense of efficacy about their teaching capabilities are better able to motivate students and enhance students' cognitive development (Ashton & Webb, 1986; Gibson & Dembo, 1984, Pintrich & Schunk, 2002). Bryant and Yan (2010) looked at the relationship between teachers' self-efficacy beliefs and adequate yearly progress (AYP) in schools in a large urban U.S. district and found that there were significant differences between the responses of teachers in schools

that met AYP and those that did not. Teachers in the schools that met AYP had much higher levels of self-efficacy in their ability to accomplish these tasks (Bryant & Yan, 2010).

Conclusion

The review of literature demonstrates that there is a clear gap between U.S. academic achievement and that of other nations on the basis of international comparative tests. While these comparisons are not ideal due to differences in context, the results of international tests have identified high-performing nations whose educational successes may offer important lessons the U.S. Canada, due to features that make it more similar to the U.S. than other high-performing nations, is identified as a nation whose education system may be reasonably contrasted with the U.S. Additionally, the review of literature demonstrates that school leadership is an essential component not only in school culture but also for schools' levels of achievement and identifies transformational leadership practices that are impactful at the level of student achievement. Whereas administrator and teacher leadership influence teacher empowerment and student achievement, this scholarship makes it clear that studying leadership and transformational leadership in particular on a comparative basis makes sense and may correlate with educational outcomes. Because teacher empowerment and satisfaction are related to leadership and educational outcomes, it is also rational to determine whether these are different cross-nationally and whether these differences relate to leadership metrics. While the review of literature is extensive, it is not exhaustive and there are many facets of international education that it does not address; likewise there are numerous other leadership strategies that this review does not cover. Despite these limitations, the combination of scholarship on

international comparative education and transformational leadership each makes it evident that there is a lack of research concerning leadership practices cross-nationally and that differences between the U.S. and Canadian educational leaders have not been adequately described. This gap in the literature paves the way for my study in which I explore teacher and administrator leadership in urban schools in U.S. and Canadian districts. The study is described in greater detail in the methodology section below.

C. METHODOLOGY

Introduction

The purpose of this study is to assess for differences in teacher and administrator leadership practices in schools from large urban districts from the United States and Canada. The leadership practices were assessed using the Leadership Practices Inventory (LPI) which assesses what Kouzes and Poser (2007) describe as five practices of exemplary leaders: modeling the way, inspiring a shared vision, challenging the process, enabling others to act, and encouraging the heart (these five practices constitute LPI subscale categories). Teachers' assessment of empowerment and satisfaction were also assessed using items from the School Participant Empowerment Scale (SPES) which assesses 6 dimensions of teacher empowerment including decision making, professional growth, status, self-efficacy, autonomy, and impact (Short & Rinehart, 1992). This section presents the study design, research hypotheses, instruments, data collection strategy, study population, and other methods involved in the study.

Research Design

This study employed surveys to study the differences between administrative and teacher leadership practices in participating schools in the U.S. and Canada. The sources of data include administrators' and teachers' assessments of their leadership behaviors as measured by the LPI instrument and teachers' perceptions of empowerment as measured by the SPES survey instrument, both of which have established validity and reliability as is described in the instrumentation section below. I also collected selected demographic information. Surveys were

performed electronically using Qualtrics software and data was tabulated in aggregate without personally identifying information. All participants were sent an e-mail asking for their voluntary participation with a link to the survey and given 4 weeks to complete the instrument. A follow-up / reminder email was sent two weeks after the initial contact and one week before the survey closed to improve response rates.

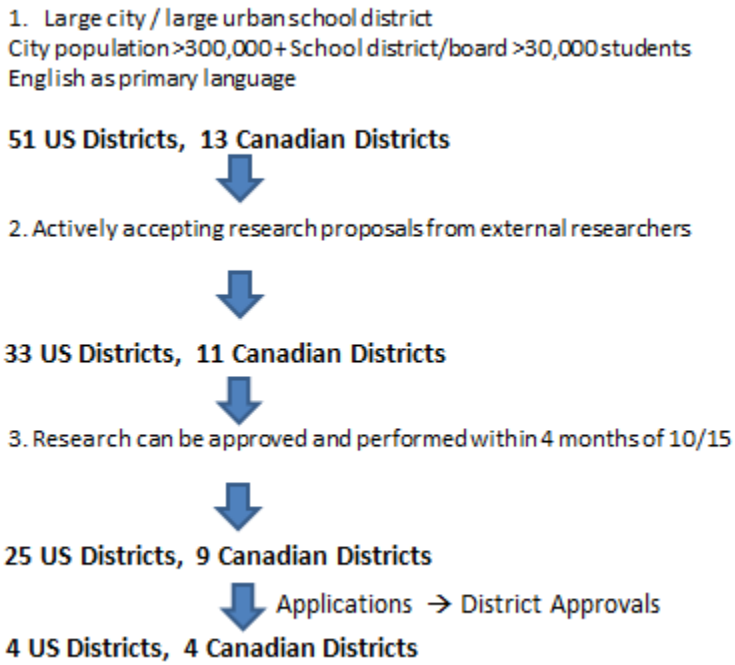
Population and Sampling

Research Participants

This work employed a purposive sample (Merriam, 1997) to capture various schools within large urban districts with varying structures and cultures. Districts were deemed eligible for study participation if they met the following criteria: city population of greater than 300,000, school district/board (the term used to describe Canadian districts) greater than 30,000 students, and English as the district's primary language. Of the 51 potentially eligible districts across the U.S. and Canada, the list was filtered by districts which were actively accepting research proposals from outside researchers and would allow research to occur within a four month timeframe. A total of twenty four districts met this criterion and research proposal applications were prepared for each of these districts. For each U.S. district and Canadian school board, formal applications were submitted and included a ten to thirty page proposal including research procedures, hypothesis, recruitment strategy, data management and disposal plan, literature review. In addition, for each district, I included verification of DePaul IRB approval (appendix 1), district specific information letters and consents, invitation letters, time requirement and estimated participant number tables, a list of preferred schools (meeting my inclusion criteria –

each school within each district individually researched & determined), recommendation letters, survey instruments, a description of benefit to participants, the district, and alignment of the study with district's strategic multi-year plan and/or research priorities. After completing the numerous revision stages of individual district external research application processes for twenty-five districts (15 U.S. and 10 Canadian), I received approval to conduct my study in four large U.S. and four large Canadian districts (referred to as districts U1-U4 and C1-C4, respectively), see figure 1 for district application process.

Figure 1. District search and application process



Within each district, schools were considered eligible for participation if they met the following inclusion criteria: high schools (students in grades 9-12) which primarily offer college preparatory programming (participating districts also called this "regular" and "standard") and

are not designated by their district to be magnet, selective enrollment, charter, gifted, special education, vocational, military, religious, single sex, adult education, or representing primarily other non-traditional curricula.

Within schools, participants were considered eligible if they were full-time employees who work in high schools that are considered to be regular programming (non-charter, magnet, religious, etc.). All participants were 21 years or older, able to read and understand English, and had access to a computer and the internet. Eligible participants were either (1) licensed teachers or (2) administrators. Administrators included both principals and vice/assistant principals.

While the initial research application was similar for all districts, modifications in methodology had to be made to accommodate district rules and therefore recruitment materials were sent in different ways (to reflect individual districts' requirements): districts U1, U2, and C2 were sent to all teachers/administrators who were found from publically available sources, while districts U3, U4, C1, C3, and C4 required that principals had to first be contacted, their permission granted, and only then they were asked to send the recruitment materials (or upon Principals' own discretion, only an invite email or whatever they thought was appropriate and were comfortable with) to distribute to their staff.

Participating Districts

Eight urban districts/school boards approved the study, four in the US and four in Canada including districts from New York, Texas, California, Ontario, Manitoba, and Alberta. All districts/boards were from cities of >300,000.

For anonymity each district's identity was not revealed to uphold district/school board confidentiality regulations. See results section for further description of districts included and appendix 2 for further details about study approval.

US Districts

District U1 is a large urban district located in the western US, city of >1 million. District U2 is an urban district located in the eastern US, in a city of over 300,000. District U3 is a large urban district located in the eastern US in a city of over 1 million. District U4 is a large urban district located in a western US city of over 1 million.

Canadian School Boards

Board C1 is a large urban district located in Western Canada in a city of over 500,000 million. Board C2 is an urban district located in Western Canada in a city of over 500,000. Board C3 is a large urban district located in Eastern Canada in a city of over 1 million. District C4 is a primarily urban district located in Eastern Canada including a city of over 300,000.

District	Country	Geography	City size	State or Province	Number of schools approved	> 50% eligible schools in district	Approval Date
U1	US	West	>1 million	State A	36	Y	12/7/2015
U2	US	East	>300,000	State B	8	Y	12/3/2015
U3	US	East	>1 million	State B	12	N	12/2/2015
U4	US	West	>1 million	State C	6	N	1/29/2016
C1	Canada	West	>500,000	Province A	12	Y	12/16/2015
C2	Canada	West	>500,000	Province B	8	Y	1/19/2016
C3	Canada	East	>1 million	Province C	25	N	1/22/2016
C4	Canada	East	>300,000	Province C	24	Y	1/26/2016

Table 1. Characteristics of districts/boards included in the study

Schools Approved In Participating Districts/School Boards

Each individual district/school board had different requirements for which schools were allowed to be included in the study. Some districts (U1, U2, C4) allowed all eligible schools to be contacted, while others (U3, U4, C1, C2) required pre-approval of the study from school administrators prior to study participation, while one district (C3) allowed me to contact a random sample of eligible schools. The processes for pre-approval differed as well with some districts approving a subset of requested schools while others required me to contact all district principals to obtain consent for school participation. After each of these processes, the number of schools approved to participate from each district were as follows: U1, thirty-six schools; U2, eight schools; U3 twelve schools; U4, six schools; C1, twelve schools; C2 eight schools; C3, twenty-five schools; C4, twenty-four schools (see table 1).

Districts were initially selected by size and were each assigned at least one "comparator" district so that any two districts with adequate response rates could be used to compare and contrast patterns of educational leadership across national settings and to reduce the risk of attributing all change events to primary differences between schools as was described by Hargreaves (2006).

As the study was initially intended to be a study of two districts, the two with the largest response rates, U1 and C1, were compared as the primary study outcome. While it would be ideal to compare each of the additional three U.S. and three Canadian districts to similar districts cross-nationally, because of low survey response rates in these districts, such comparisons would have very limited statistical power. Therefore, to ensure that the surveys completed by the six additional districts were included in the data, I performed an analysis of all participants from

U.S. districts versus all participants in Canadian districts as a secondary outcome. While heterogeneity across districts may complicate this analysis, the comparison of four districts from each country is also of interest both in that it provides a larger cross-section of urban schools from each country and is therefore more representative of the country's education system than a single district and because this analysis looks at a larger overall sample size and is therefore more statistically powerful.

Instrumentation

Demographic Information

Both teachers and administrators were asked demographic information including their school (selected from list), their gender, ethnicity, age (by strata), level of education, teaching experience (by strata), and years at current school (by strata). Participants had no way to include their name on the survey and no IP tracking software (which was purposefully disabled) was used to ensure the survey contained no personal identifiers linking a survey with a respondent. The survey used school codes to differentiate the schools participating in this study, but these were also de-identified to prevent any linking of data to respondents. The demographic information collected and strata used are detailed in appendix 3.

Survey Instruments

A. Administrator Survey (see appendix 4)

Administrators received a complete version of the Kouzes and Posner Leadership Practices Inventory (LPI) self-assessment that consists of 30 questions which fall into the 5 leadership categories. A copy of the full administrator survey is attached as an appendix.

B. Teacher Survey (see appendix 5)

Teachers received surveys with components from the LPI with text modified to reflect their leadership in the classroom (12 items derived from 4 of the 5 leadership subscale categories) and from the School Participant Empowerment Scale (SPES) which assess 6 aspects of teacher empowerment. I assessed teacher satisfaction using four previously validated questions. A copy of the full teacher survey is attached as an appendix.

LPI Validation and Use

The Learning Practices Inventory survey instrument for understanding leadership practices has been used extensively; data from over 350,000 respondents has helped demonstrate the validity of the LPI instrument (Kouzes & Posner, 2002). The 30 questions and 5 subscales of the LPI are measured on a Likert scale in which participants rate themselves from one to ten based on the extent to which they engage in each specified behavior. The LPI questions each have internal reliability coefficients above the 0.75 level as measured by Cronbach's alpha. Furthermore, this reliability for LPI subscales is consistent across different participant groups (including educators, but also business, healthcare, and religious sectors), and amongst various demographic groups including race, nationality, gender, and marital status. In the education sphere, this instrument has been widely used to assess administrative leadership practices.

SPES Validation and Use

First described by Short and Rinehart in 1992, the SPES is a 38-item instrument score on a 5-point Likert scale (scored from 1=strongly disagree to 5=strongly agree for each statement).

Factor analysis of the SPES revealed six dimensions each of which demonstrated a Cronbach's alpha of 0.80 or greater with the overall scale reliability of 0.94. Since its inception, the SPES has been used in numerous educational studies to demonstrate that teacher efficacy correlates with other relevant outcomes including job satisfaction, participation in decision-making, instructional practice, and student academic achievements (Bogler, 2004).

Sample Size And Power Calculation

Statistical power represents the likelihood of a statistical test detecting an effect if the effect actually exists. In districts U1 and C1, there were 161 and 85 teacher responses from U1 and C1 respectively and 20 administrator responses from each district. With LPI and SPES subscale average differences of 0.75and standard deviations of 0.5, this gives me >80% power to detect statistical significance at the alpha=0.05 level with an effect size of 1.25 among teachers but only 48% power to detect difference among administrators. In the cross-national study with 213 US teachers and 163 Canadian teachers and 44 US administrators and 40 Canadian administrators, I will have >80% power amongst both teachers and administrators. Therefore, I have sufficient (>80%) confidence and differences seen in the cross-national comparisons or cross-district teacher comparison are real, but there remains a ~50% chance that I will fail to detect administrator differences across districts even if they exist.

Data Collection, Inclusion, and Analysis Procedure

All de-identified survey data were collected using DePaul University's online approved Qualtrics software and were exported as tables into Microsoft Excel where data were combined,

sorted, and cleaned. Incomplete surveys and surveys with no variability between all responses were excluded from the analysis. Surveys missing only demographic data were included if they indicated the participants' school. Data were then exported and coded into IBM SPSS 23 software (IBM Corp, Armonk, NY, 2012) where they were analyzed using rigorous statistical methodologies.

Research Questions

In districts U1 and C1 (the two districts with the largest response rates):

What *differences* exist in leadership practices between:

1. Administrators (principals and vice principals)?
2. Teachers?

What *relationships* exist between:

3. Administrative and teacher leadership practices?
4. Administrative leadership practices and teacher empowerment?

Statistical Testing: Hypotheses and Assumptions

Inferential Statistics, Null Hypotheses, and Tests Of Statistical Significance

Studies which seek to answer questions about populations based on data collected from samples of those populations use inferential statistics which utilize statistical procedures for "making inferences about characteristics of populations based on data collected from samples that were selected to represent those populations" (Gall, Gall, & Borg, 2010, 197). In inferential

statistics, when two samples are compared, the procedure used to compare whether there are differences in means or variances is based on testing of two possible inferences, the null hypothesis and the alternative hypothesis (Gall, Gall, & Borg, 2010). The null hypothesis "states that no effect will occur, or that no differences or relationships will be found (Charles & Mertler, 2002, 62) and more specifically "makes a prediction that in the general population, no relationship or no significant difference exists between groups on a variable. The wording is 'There is no difference (or relationship) between the groups' " (Cresswell, 2009, p 134). In this type of statistics, the null hypothesis is set to the 'no/none' standard and then tests of statistical significance are used to assess the likelihood that one can use data to "find contradictions" and reject the null hypothesis (Charles & Mertler, 2002). Statistical hypothesis testing utilizes tests (in this study Mann-Whitney U tests and Spearman's rho) in which one assesses the probability that the null hypothesis is true based on the data by assessing a p-value. A p-value refers to the "percentage of occasions that a chance difference between mean scores of a certain magnitude will occur [by random chance] when the means are identical" (Gall, Gall & Borg, 2010, p 197). It is therefore imperative that the null hypothesis be a statement of no difference so that the statistical tests can assess how likely (based on the p-value) the data would approximate a distribution that differs from the null hypothesis.

Statistical Questions

1a. What *differences* exist in leadership practices between administrators in the urban secondary school districts U1 and C1?

Null hypothesis: There will be no difference in LPI scores between administrators between the district U1 and district C1 administrators

H_0: $\mu_{U1.Admin\ LPI} = \mu_{C1\ Admin\ LPI}$

Alternative hypothesis: There will be differences in LPI scores between administrators between the U1 and C1 administrators

$H_{1:}\ \mu_{U1\ Admin\ LSI} \neq \mu_{C1\ Admin\ LPI}$

1b. What *differences* exist in leadership practices between US and Canadian administrators across all participant districts?

Null hypothesis: There will be no difference in LPI scores between administrators between US and Canadian administrators

H_0: $\mu_{U.S.Admin\ LPI} = \mu_{Canada\ Admin\ LPI}$

Alternative hypothesis: There will be differences in LPI scores between administrators between US and Canadian administrators

H_1: $_{U.S\ Admin\ LPI} \neq \mu_{Canada\ Admin\ LPI}$

2a. What *differences* exist in leadership practices (LPI), satisfaction, and empowerment (SPES) between teachers in districts U1 and C1?

Null hypothesis: There will be no difference in LPI scores, satisfaction, or SPES score among teachers of the U1 and C1 districts

H_0: $\mu_{U1Teacher\ LPI} = \mu_{C1\ Teacher\ LPI}$

Alternative hypothesis: There will be differences in LPI, satisfaction, and SPES scores among teachers of the U1 and C1 districts.

H_1: $\mu_{U1Teacher\ LPI} \neq \mu_{C1\ Teacher\ LPI}$

2b. What *differences* exist in leadership practices (LPI), satisfaction, and empowerment (SPES) between teachers in all participant US and Canadian districts?

Null hypothesis: There will be no difference in LPI scores, satisfaction, or SPES score between US and Canadian teachers

H_0: $\mu_{U.S.\ Teacher\ LPI} = \mu_{Canada\ Teacher\ LPI}$

Alternative hypothesis: There will be difference in LPI scores, satisfaction, or SPES score between US and Canadian teachers

H_1: $\mu_{U.S\ Teacher\ LPI} \neq \mu_{Canada\ Teacher\ LPI}$

3. What *relationships* exist between administrative and teacher leadership practices?

Null hypothesis: Teachers and administrators from the same school will have LPI scores which do not correlate with each other (spearman rho = 0)

H_0: No correlation between teacher and administrator LPI scores; (spearman R = 0)

Alternative hypothesis: Teachers and administrators from the same school will have LPI scores which do significantly correlate with each other (spearman rho does not equal 0)

H_1: Correlation between teacher and administrator LPI scores; (spearman $R \neq 0$)

4. What *relationships* exist between administrative leadership practices and teacher empowerment?

Null hypothesis: There will be no correlation between administrator LPI scores and teacher SPES scores at the same school; (spearman rho = 0)

H_0: No correlation between administrator LPI score and teacher SPES score; (spearman $R = 0$)

Alternative hypothesis: There will be a significant correlation between administrator LPI scores and teacher SPES scores at the same school; (spearman rho does not equal to 0)

H_1: There is a correlation between administrator LPI score and teacher SPES score; (spearman $R \neq 0$)

Statistical Analysis

Hypothesis Testing

Questions 1 and 2:

To assess the null hypothesis that the mean LPI score between administrators (q1) or teachers (q2) in the two countries will be the same, because results were non-normally distributed, non-parametric testing needed to be performed. The Mann-Whitney test is useful in comparing the average score across two groups in which the assumption of normality cannot be

made; compared to a 2 sample student's t-test, the Mann Whitney test can be used with ordinal data, is more robust to outliers, and is more efficient. I used a Mann-Whitney U test to determine if there are differences between the average LPI and SPES scores between U.S. and Canadian administrators. Additional Mann-Whitney tests were used to assess each of the 5 LPI subscale scores (Modeling the Way, Inspiring a Shared Vision, Challenging the Process, Enabling Others to Act, and Encouraging the Heart) between (1) administrators and (2) teachers as well as teacher SPES subscale scores (decision making, professional growth, status, self-efficacy, autonomy, and impact) and teacher satisfaction.

Questions 3 and 4:

To determine if relationships exist between the administrators' LPI scores and either (q3) teachers' LPI scores or (q4) teachers' SPES scores, correlation analysis was performed and Spearman's rho correlations were computed for all observations. The Spearman test is appropriate because it is a nonparametric measure of statistical dependence between two variables (administrator LPI score with teacher LPI or SPES scores in this study). It assesses how well the relationship between two variables can be described using a monotonic function and therefore is valid for non-linear correlations (unlike Pearson's correlation which assumes normality and linear correlations).

Statistical Testing Assumptions

In performing the present study, I need to make a number of statistical assumptions:

1. For statistical testing in questions 1 and 2the Mann- test assumes that there is an underlying continuity from low to high in the dependent variable before ranking and that

observations are independent. In other words, these tests assume that the 1-5 Likert scales used by the LPI and SPES surveys represent five levels in which each level is distinctly different and can be distinguished by order (i.e. strongly agree > agree > neutral > disagree > strongly disagree) and that each individual's responses are not impacted by the responses of others.

2. For questions 3 and 4, to perform Spearman's rho, it is assumed that the variables are at least ordinal and that scores of each variable are monotonically related to the other variable. For this study, this means that the tests assume that there is a discernable relationship between administrator LPI scores and teachers' (1) LPI scores or (2) SPES scores which are not influenced by other factors.

Additional Non-Statistical Assumptions:

1. My personal beliefs and biases did not impact on the outcome of the interpretation of the statistical outcomes of the study.
2. The LPI and SPES remained valid in the study populations in which I employed them
3. Minor modifications that I made to the LPI questions (to reflect the teachers as participants) did not change their intrinsic meaning.
4. The validated LPI and SPES survey questions used remained valid although the teachers survey did not contain the entirety of either of these instruments.
5. There are no major cultural or linguistic differences in participants' understanding of the questions that led to different interpretations between participants in the U.S. and Canada.

D. RESULTS

Survey Responses By Type, District, Country, And Demographics.

624 individual began the survey and 460 individuals completed the survey instrument and were included in the final analysis. Of these, 376 were teachers and 84 were administrators, 257 were from US districts and 203 were from Canadian school districts. Breakdown of response by district is and by country is shown in table 2. Districts beginning with U are US districts while those beginning with C are Canadian districts. District U1 (181 responses) and district C1 (105 responses) were selected as the primary districts for analysis given their higher response numbers. Histograms representing the responses by demographic variables are shown for districts U1 vs C1 (figure 2a-2e) and across the 4 US and 4 Canadian districts considered together (figure 3a-3e). The most notable trends include that US participants were more likely to have a master's degree while Canadian teachers were more likely to have a bachelor's and that US respondents were more likely to have 0-10 years of experience compared to Canadians who were more likely to have 16 or more years' experience. Canadian districts were more Caucasian that US districts, but gender and age were distributed similarly. The 3 US states which participating districts were drawn from included New York, Texas, and California while the three Canadian provinces with participant districts included Ontario, Alberta, and Manitoba. For confidentiality guaranteed by the individual districts/boards, the individual cities represented have been de-identified with the U/C 1-4 classifications.

Table 2. Survey Respondents

A.		Teacher	Administrator	Total
District	**U1**	**161**	**20**	**181**
	U2	46	3	49
	U3	6	18	24
	U4	0	3	3
	C1	**85**	**20**	**105**
	C2	47	5	52
	C3	24	5	29
	C4	7	10	17
Total		376	84	460

Table 2a. Survey respondents: Teachers and administrators listed by district

B.		Teacher	Administrator	Total
Country	US	213	44	257
	Canada	163	40	203
Total		376	84	460

Table 2b. Survey respondents: Teachers and administrators listed by country

Figures 2 and 3. Graphs of demographic variables, split by district U1 vs C1 (Figure 2 a-e), and US vs Canada (Figure 3 a-e)

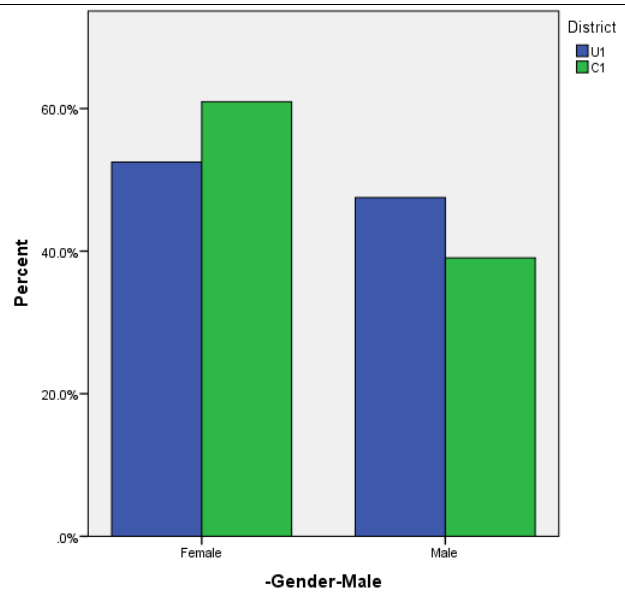


Figure 2a. Gender, District level

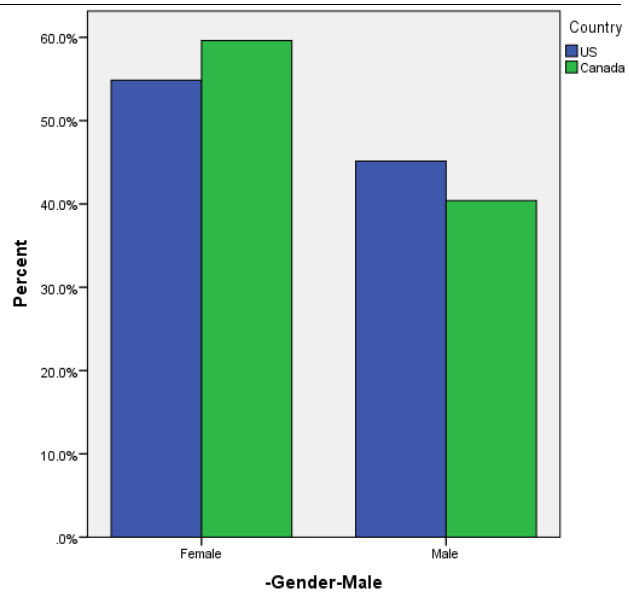


Figure 3a. Gender, Country level

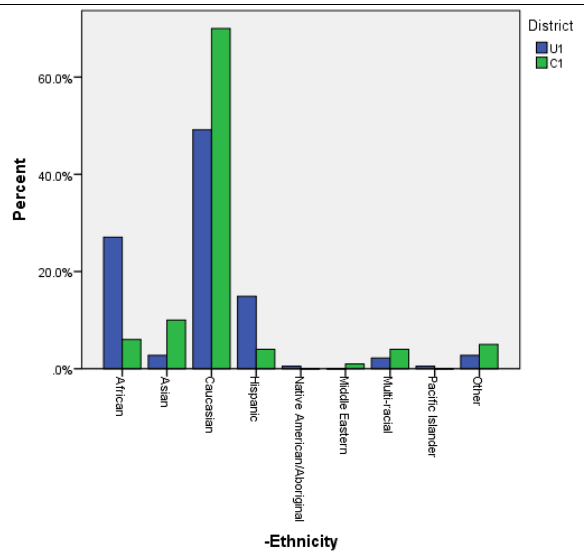


Figure 2b. Ethnicity, District level

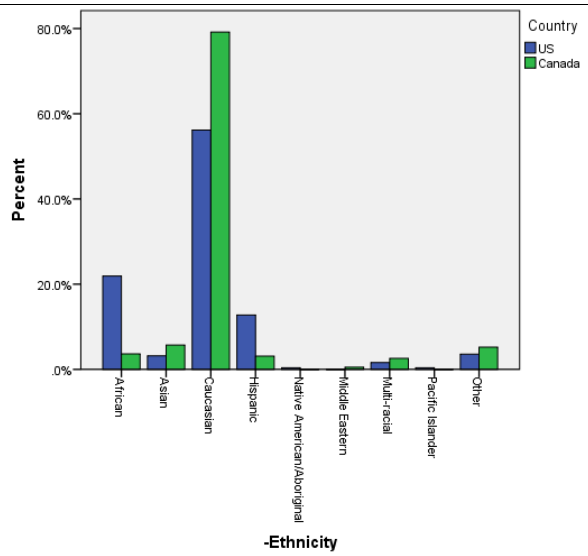


Figure 3b. Ethnicity, Country level

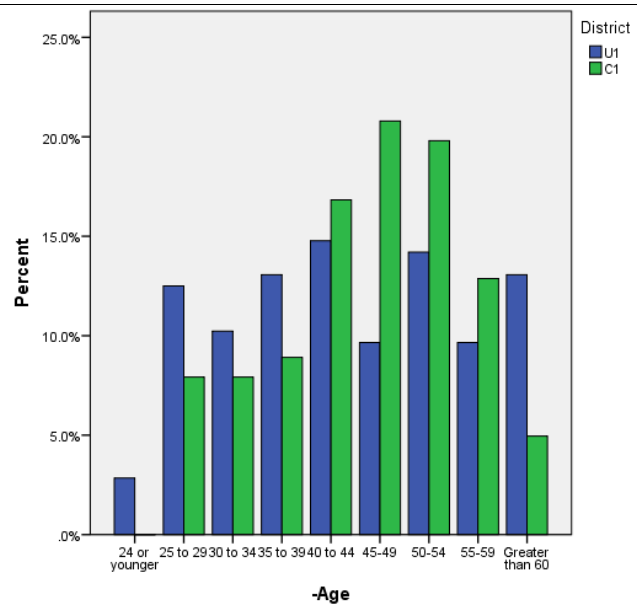


Figure 2c. Age strata, District level

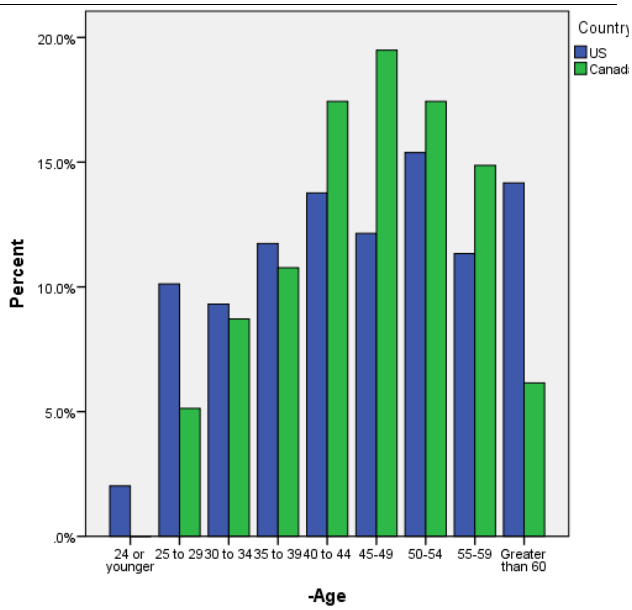


Figure 3c. Ethnicity, Country level

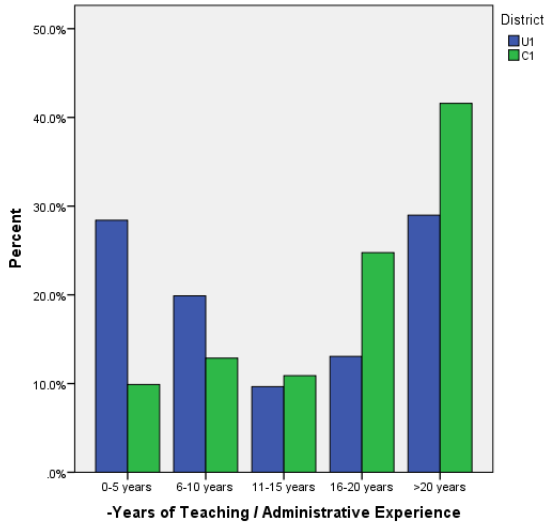


Figure 2d. Years of experience, District level

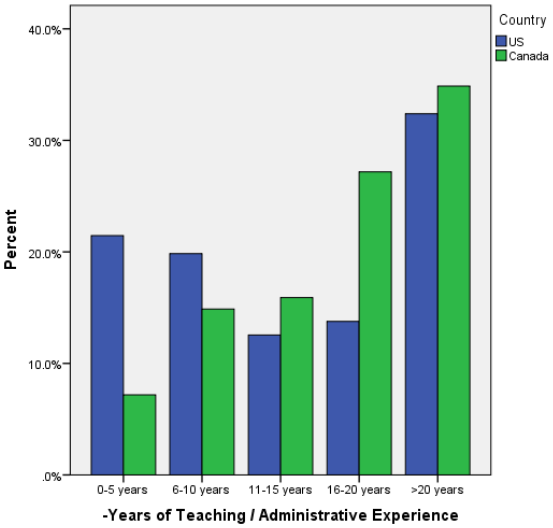


Figure 3d. Years of experience, Country level

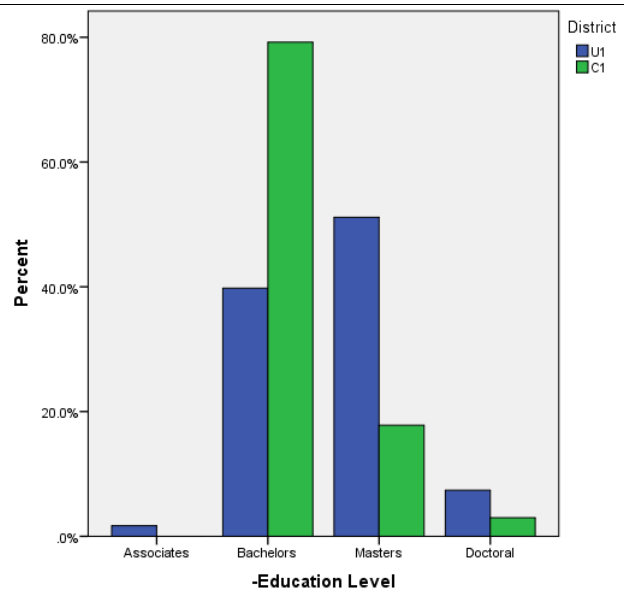


Figure 2e. Education level, District level

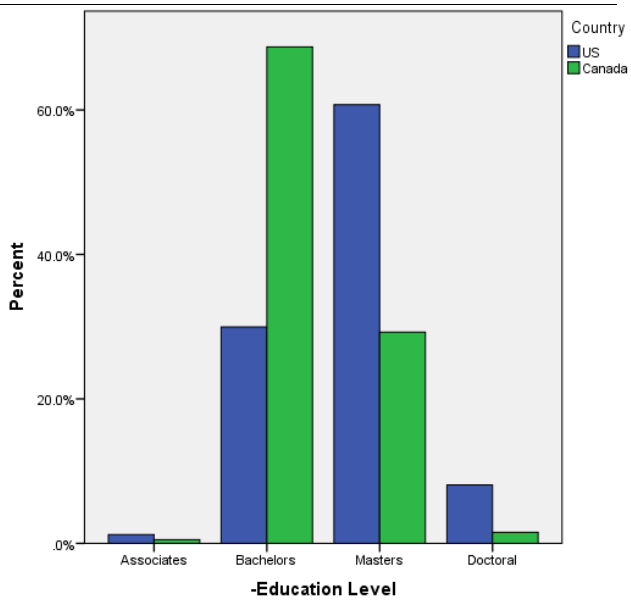


Figure 3e. Education level, Country level

Question 1: Administrator Leadership Practices

1a. What differences exist in leadership practices between administrators in the urban secondary school districts U1 and C1?
1b. What differences exist in leadership practices between U.S. and Canadian administrators across all participant districts?

Findings for question 1a: Differences between administrator leadership across districts U1 and C1

In order to determine whether there were any differences in administrative leadership between the two districts, I first assessed the means and standard deviations for each of the 5 LPI subscale sections. Results are shown in the descriptive statistics table (Table 3). A Kolmogorov-Smirnov test was then used to assess the normality of the distributions for each category (Table 4). Because all categories demonstrated significant ($p<0.05$) deviation from a normal distribution, non-parametric tests needed to be used to assess differences between districts. In order to assess differences in LPI scores, I performed a Mann-Whitney U test comparing the ranks of administrators' LPI sub-score between the two districts (20 administrators from each district). The ranks table (Table 5) shows the mean rank and sum of ranks for each sub-scale category. To determine if the differences in ranks observed were significant, the 2-tailed p-values calculated by the Mann-Whitney test are listed in the 4th row of the test statistics table (Table 6). While some differences were apparent when looking at averages and rank averages, particularly for "challenge the process" and "enable others to act", none of the 5 sub-scale categories or the overall LPI scores demonstrated a difference at a p-value of $p < 0.05$ which indicates that *there are no statistically significant differences between administrators' self-reported leadership behaviors between the two districts*. Based in this result, I accept the null hypothesis that there are no significant differences between LPI scores between districts U1 and C1. To further analyze the 30 individual LPI questions, I also performed Mann-Whitney tests

for each of the 30 individual LPI items (Supplemental table 1). Only 2 of the 30 items ("I experiment and take risks even when there is a chance of failure" and "I praise teachers for a job well done") showed statistically significant differences with district C1 administrators reporting these practices far more frequently than administrators from district U1 (mean ranks of 24.4 (C1) vs 16.6 (U1) and 24.5 (C1) vs 16.5 (U1) for the two questions respectively).

Table 3. Descriptive statistics for districts U1 and C1 administrator LPI scores. In this table mean, standard deviation, and range averages are shown for each of the LPI subscale categories in districts U1 and C1

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Model the way P	U1	20	14.49	0.36	0.08	14.32	14.66	13.67	15.00
	C1	20	14.58	0.32	0.07	14.43	14.73	14.00	15.00
Challenge the process admin	U1	20	14.27	0.46	0.10	14.05	14.48	13.17	15.00
	C1	20	14.50	0.35	0.08	14.34	14.66	13.83	15.00
Enable others to act admin	U1	20	14.40	0.36	0.08	14.23	14.57	13.67	15.00
	C1	20	14.58	0.31	0.07	14.44	14.73	14.00	15.00
Encourage the heart admin	U1	20	14.20	0.54	0.12	13.95	14.45	13.50	15.00
	C1	20	14.43	0.48	0.11	14.20	14.65	13.50	15.00
Inspire a shared vision admin	U1	20	14.13	0.73	0.16	13.79	14.47	12.00	15.00
	C1	20	14.19	0.57	0.13	13.92	14.46	13.00	15.00

Table 4. Kolmogorov-Smirnov test to assess normality of distributions. In this table, each of the LPI subscale categories is compared to a normal distribution using the one-sample Kolmogorov-Smirnov test. All of the subscales show significant ($p<0.05$) deviation from the normal distribution indicating that parametric tests which assume a normal distribution (t-tests, ANOVA, etc.) are not advised and that instead non-parametric tests (Mann-Whitney, Kruskal-Wallis, etc.) be used.

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Model the way P is normal with mean 14.47 and standard deviation 0.416.	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
2	The distribution of Challenge the process admin is normal with mean 14.32 and standard deviation 0.516	One-Sample Kolmogorov-Smirnov Test	.001[1]	Reject the null hypothesis.
3	The distribution of Enable others to act admin is normal with mean 14.46 and standard deviation 0.359	One-Sample Kolmogorov-Smirnov Test	.005[1]	Reject the null hypothesis.
4	The distribution of Encourage the heart admin is normal with mean 14.28 and standard deviation 0.550	One-Sample Kolmogorov-Smirnov Test	.007[1]	Reject the null hypothesis.
5	The distribution of Inspire a shared vision admin is normal with mean 14.13 and standard deviation 0.695	One-Sample Kolmogorov-Smirnov Test	.002[1]	Reject the null hypothesis.
6	The distribution of adminLPI is normal with mean 14.33 and standard deviation 0.434.	One-Sample Kolmogorov-Smirnov Test	.021[1]	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[1]Lilliefors Corrected

Table 5. Administrator LPI ranks for districts U1 and C1. Ranks across the 40 administrators in the two districts were assigned for all administrators' responses for each LPI subscale category, this table shows both the average (mean) rank for each district as well as the sum of ranks for each district.

	District	N	Mean Rank	Sum of Ranks
Model the way P	U1	20	19.13	382.50
	C1	20	21.88	437.50
	Total	40		
Challenge the process admin	U1	20	17.45	349.00
	C1	20	23.55	471.00
	Total	40		
Enable others to act admin	U1	20	17.53	350.50
	C1	20	23.48	469.50
	Total	40		
Encourage the heart admin	U1	20	18.13	362.50
	C1	20	22.88	457.50
	Total	40		
Inspire a shared vision admin	U1	20	20.38	407.50
	C1	20	20.63	412.50
	Total	40		
Admin LPI avg	U1	20	18.25	365.00
	C1	20	22.75	455.00
	Total	40		

Table 6. Test statistics for Mann-Whitney test between districts U1 and district C1 administrator LPI scores. This table shows the test statistics for the Mann-Whitney test performed across districts U1 and C1. The statistical significance is represented by the Asymp. Sig. (2-tailed) row and statistically significant results would be those with values of $p<0.05$. Because all 5 subscales and the overall LPI score are all >0.05, none is deemed statistically significant and I accept the null hypothesis that there are no significant differences across administrators' LPI scores in districts U1 and C1.

	Model the way P	Challenge the process admin	Enable others to act admin	Encourage the heart admin	Inspire a shared vision admin	Admin LPI avg
Mann-Whitney U	172.500	139.000	140.500	152.500	197.500	155.000
Wilcoxon W	382.500	349.000	350.500	362.500	407.500	365.000
Z	-.760	-1.670	-1.625	-1.297	-.068	-1.219
Asymp. Sig. (2-tailed)	.447	.095	.104	.195	.946	.223
Exact Sig. [2*(1-tailed Sig.)]	.461[b]	.102[b]	.108[b]	.201[b]	.947[b]	.231[b]

a. Grouping Variable: District

b. Not corrected for ties.

Findings For Question 1b: Differences Between US And Canadian Administrators Across All Participant Districts

In order to determine whether there were any differences in administrative leadership between all participant districts across the U.S. and Canada, I again assessed the means and standard deviations for each of the 5 LPI subscale sections. Results are shown in the descriptive statistics table (Table 7). A Kolmogorov-Smirnov test again demonstrated significant ($p<0.05$) deviation from a normal distribution (data not shown), and thus again non-parametric tests were used to assess differences between administrators from the 2 countries. In order to assess differences in LPI scores, I performed a Mann-Whitney U test comparing the ranks of administrators' LPI sub-scores between the all US (44) and all Canadian (40) administrator respondents. The ranks table (Table 8) shows the mean rank and sum of ranks for each sub-scale category. To determine if the differences in ranks observed were significant, the 2-tailed p-values calculated by the Mann-Whitney test are listed in the test statistics table (Table 9). Much as there were no differences across the two largest districts, none of the 5 sub-scale categories or the overall LPI scores demonstrated difference at a p-value of $p < 0.05$ which indicates that *there were no statistically significant differences between administrators self-reported leadership behaviors between participant districts' administrators in the two countries.* This means that I accept the null hypothesis that there are no significant differences between LPI scores between U.S. and Canadian administrators.

To assess if there were differences in any of the 30 individual LPI questions, I performed Mann-Whitney tests for each of the 30 individual LPI items (Supplemental table 2). In this case, 3 of the 30 items ("I follow through on the promises and commitments that I make", "I ask for

feedback on how my actions affect teachers' performance", and "I praise teachers for a job well done") showed statistically significant differences with U.S. administrators reporting asking for feedback more frequently (mean rank 47.19) than Canadian administrators (mean rank 37.34) while Canadian administrators reported following through on commitments and praising teachers (mean ranks 47.45 and 48.88) significantly more than their U.S. counterparts (mean ranks 38.0 and 36.70).

Table 7. Descriptive statistics for administrator LPI in 8 districts surveyed. In this table mean, standard deviation, and range averages are shown for each of the LPI subscale categories in all US (U1,U2,U3,U4) and Canadian (C1,C2,C3,C4) districts.

						95% Confidence Interval for Mean			
		N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
Model the way P	U1	20.00	14.49	0.36	0.08	14.32	14.66	13.67	15.00
	U2	3.00	14.28	0.98	0.56	11.85	16.70	13.17	15.00
	U3	18.00	14.39	0.42	0.10	14.18	14.60	13.50	15.00
	U4	3.00	14.83	0.29	0.17	14.12	15.55	14.50	15.00
	C1	20.00	14.58	0.32	0.07	14.43	14.73	14.00	15.00
	C2	5.00	14.20	0.66	0.30	13.38	15.02	13.67	15.00
	C3	5.00	14.53	0.14	0.06	14.36	14.71	14.33	14.67
	C4	10.00	14.38	0.42	0.13	14.09	14.68	13.50	14.67

	Total	84.00	14.47	0.42	0.05	14.38	14.56	13.17	15.00
Challenge the process admin	U1	20.00	14.27	0.46	0.10	14.05	14.48	13.17	15.00
	U2	3.00	13.72	1.27	0.73	10.56	16.88	12.33	14.83
	U3	18.00	14.40	0.43	0.10	14.18	14.61	13.50	15.00
	U4	3.00	14.78	0.19	0.11	14.30	15.26	14.67	15.00
	C1	20.00	14.50	0.35	0.08	14.34	14.66	13.83	15.00
	C2	5.00	13.90	0.83	0.37	12.87	14.93	13.00	15.00
	C3	5.00	14.43	0.35	0.15	14.00	14.86	14.00	14.83
	C4	10.00	14.15	0.49	0.16	13.80	14.50	13.33	15.00
	Total	84.00	14.32	0.52	0.06	14.21	14.44	12.33	15.00
Enable others to act admin	U1	20.00	14.40	0.36	0.08	14.23	14.57	13.67	15.00
	U2	3.00	14.22	0.51	0.29	12.96	15.49	13.67	14.67
	U3	18.00	14.44	0.35	0.08	14.26	14.61	13.50	14.83
	U4	3.00	14.78	0.25	0.15	14.15	15.41	14.50	15.00
	C1	20.00	14.58	0.31	0.07	14.44	14.73	14.00	15.00
	C2	5.00	14.33	0.68	0.30	13.49	15.17	13.50	15.00
	C3	5.00	14.43	0.09	0.04	14.32	14.55	14.33	14.50
	C4	10.00	14.32	0.25	0.08	14.14	14.50	14.00	14.67

	Total	84.00	14.45	0.36	0.04	14.37	14.52	13.50	15.00
Encourage the heart admin	U1	20.00	14.20	0.54	0.12	13.95	14.45	13.50	15.00
	U2	3.00	13.83	1.04	0.60	11.25	16.42	13.00	15.00
	U3	18.00	14.33	0.57	0.14	14.05	14.62	13.33	15.00
	U4	3.00	14.72	0.35	0.20	13.86	15.58	14.33	15.00
	C1	20.00	14.43	0.48	0.11	14.20	14.65	13.50	15.00
	C2	5.00	14.03	0.83	0.37	13.00	15.06	13.17	15.00
	C3	5.00	14.43	0.09	0.04	14.32	14.55	14.33	14.50
	C4	10.00	14.10	0.45	0.14	13.78	14.42	13.17	15.00
	Total	84.00	14.28	0.55	0.06	14.16	14.40	13.00	15.00
Inspire a shared vision admin	U1	20.00	14.13	0.73	0.16	13.79	14.47	12.00	15.00
	U2	3.00	13.28	1.51	0.87	9.52	17.03	12.17	15.00
	U3	18.00	14.17	0.63	0.15	13.85	14.48	13.00	15.00
	U4	3.00	14.56	0.51	0.29	13.29	15.82	14.00	15.00
	C1	20.00	14.19	0.57	0.13	13.92	14.46	13.00	15.00
	C2	5.00	14.13	0.84	0.38	13.08	15.18	13.17	15.00
	C3	5.00	14.40	0.58	0.26	13.67	15.13	13.67	15.00
	C4	10.00	13.92	0.67	0.21	13.44	14.39	12.83	15.00

Total	84.00	14.13	0.69	0.08	13.98	14.28	12.00	15.00

Table 8. Administrator LPI ranks across 8 participant districts. Ranks across the 84 administrators from all 8 participant districts in the two countries were assigned for all administrators' responses for each LPI subscale category, this table shows both the average (mean) rank for each country as well as the sum of ranks for each country.

	Country	N	Mean Rank	Sum of Ranks
Model the way P	US	44	42.23	1858.00
	Canada	40	42.80	1712.00
	Total	84		
Challenge the process admin	US	44	42.41	1866.00
	Canada	40	42.60	1704.00
	Total	84		
Enable others to act admin	US	44	41.56	1828.50
	Canada	40	43.54	1741.50
	Total	84		
Encourage the heart admin	US	44	42.13	1853.50
	Canada	40	42.91	1716.50
	Total	84		
Inspire a shared vision admin	US	44	42.77	1882.00
	Canada	40	42.20	1688.00
	Total	84		
Admin LPI avg	US	44	42.36	1864.00
	Canada	40	42.65	1706.00
	Total	84		

Table 9. Mann-Whitney test statistics for administrator LPI cross-nationally. This table shows the test statistics for the Mann-Whitney test performed across all US and Canadian participant districts. The statistical significance is represented by the Asymp. Sig. (2-tailed) row and statistically significant results would be those with values of $p<0.05$. Because all 5 subscales and the overall LPI score are all >0.05, none is deemed statistically significant and I accept the null hypothesis that there are no significant differences across administrators' LPI scores between US and Canadian respondents.

	Model the way	Challenge the process	Enable others to act	Encourage the heart	Inspire a shared vision	LPI average
Mann-Whitney U	868.000	876.000	838.500	863.500	868.000	874.000
Wilcoxon W	1858.000	1866.000	1828.500	1853.500	1688.000	1864.000
Z	-.109	-.036	-.376	-.149	-.108	-.054
Asymp. Sig. (2-tailed)	.913	.971	.707	.882	.914	.957

a. Grouping Variable: Country

Question 2. Teacher Leadership Practices, Satisfaction, And Empowerment

2a. What differences exist in leadership practices (LPI), satisfaction, and empowerment (SPES) between teachers in districts U1 and C1?
2b. What differences exist in leadership practices (LPI), satisfaction, and empowerment (SPES) between teachers in all participant U.S. and Canadian districts?

Findings for question 2a: Differences between teacher leadership, satisfaction, and empowerment across districts U1 and C1

In order to determine whether there were any differences in teacher leadership, satisfaction, and empowerment between the two districts, I first assessed the means and standard deviations for the relevant 4 LPI subscale sections, satisfaction, and the six SPES subscale sections (results are shown in Table 10). A Kolmogorov-Smirnov test was then used to assess the normality of the distributions for each category. All categories demonstrated significant

($p<0.001$) deviation from a normal distribution, and therefore non-parametric tests need to be used to assess differences between districts (Table 11). In order to assess differences in LPI/satisfaction/SPES scores, I performed a Mann-Whitney U test comparing the ranks of teachers' section scores between the two districts (161 teachers from district U1 and 85 from district C1). The ranks table (Table 12) shows the mean rank and sum of ranks for each sub-scale category. To determine if the differences in ranks observed were significant, the 2-tailed p-values calculated by the Mann-Whitney test were calculated and are listed in the test statistics table (Table 13). Significant differences were seen in 1 of the 4 LPI subscale categories, "challenge the process", and U1 teachers had significantly higher scores than C1 teachers (mean rank of 132.07 vs 107.27, $p=0.008$). However, there were no differences seen in the overall LPI scores. Satisfaction scores (the average of 4 questions about teachers' satisfaction with their job and school leadership team) were significantly higher in district C1 than in district U1 (mean rank of 156.01 vs 106.34, $p<0.001$). The overall SPES score was significantly higher in district C1 than in U1 (mean rank 142.68 vs 114.20, $p=0.003$) and the teachers from this district also showed significantly higher scores in 3 of the 6 SPES sub-scale categories, "professional growth" ($p<0.001$), "autonomy" ($p=0.034$) and "decision-making" ($p=0.035$). *Given these differences, there are statistically significant differences between teachers' satisfaction and empowerment between the two districts.* While there are not overall differences in leadership practices, U1 teachers demonstrated higher levels of challenging the process. Given these results, I reject the null hypothesis that there are no significant differences between LPI, satisfaction, and SPES scores between districts U1 and C1.

To assess if there were differences in any of the 30 individual questions, I also performed Mann-Whitney tests for each of the 30 individual items administered to the teachers (12 representing the 4 LPI subscale categories, 4 representing satisfaction, and 12 representing the 6 SPES categories (Supplemental table 3). 11 of the 30 questions showed significant differences with U1 having higher scores on 3 LPI questions ("I challenge students to try out new and innovative ways to do their work", "I search outside formal boundaries for innovative ways to improve what I do", "I make it a point to let students know about my confidence in their abilities"), C1 had higher scores on 1question ("I treat students with dignity and respect"), and district C1 had higher scores on all 4 satisfaction questions ("I am satisfied with my job", "I am satisfied with my teaching environment", "I am satisfied with the current leadership team", "I am satisfied with the type of leaders we"), and 3 SPES questions ("I work at a school where students come first", "I have the support and respect of my colleagues", "I am able to teach as I choose"). For further statistical assessment of each question, see supplemental table 3.

Table 10. Descriptive statistics for teacher LPI and SPES subscales across districts U1 and C1. In this table mean, standard deviation, and range averages are shown for each of the LPI, satisfaction, and SPES subscale categories in districts U1 and C1

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Model the way teachers	U1	161	14.449	0.46884	0.03695	14.376	14.522	13	15
	C1	85	14.416	0.56789	0.0616	14.293	14.538	11	15
Challenge the process teachers	U1	161	14.364	0.59489	0.04688	14.272	14.457	12.67	15
	C1	85	14.118	0.74483	0.08079	13.957	14.278	11	15
Enable others to act teachers	U1	161	14.64	0.43297	0.03412	14.572	14.707	13	15
	C1	85	14.612	0.51137	0.05547	14.502	14.722	11	15
Encourage the heart teachers	U1	161	14.559	0.53308	0.04201	14.476	14.642	13	15
	C1	85	14.443	0.69463	0.07534	14.293	14.593	11	15
Satisfaction teachers	U1	161	8.3789	1.04474	0.08234	8.2163	8.5415	6	10
	C1	85	9.0824	0.78337	0.08497	8.9134	9.2513	6.75	10
Decision-making teachers	U1	161	7.9752	0.97836	0.07711	7.8229	8.1274	6	10
	C1	85	8.2382	0.82907	0.08993	8.0594	8.4171	6.5	10
Professional growth teachers	U1	161	8.8851	0.907	0.07148	8.7439	9.0263	6	10
	C1	85	9.3176	0.69799	0.07571	9.1671	9.4682	7	10
Status teachers	U1	161	9.3509	0.63454	0.05001	9.2522	9.4497	6	10
	C1	85	9.5	0.49401	0.05358	9.3934	9.6066	7.5	10
Self-efficacy teachers	U1	161	9.323	0.77944	0.06143	9.2017	9.4443	6	10
	C1	85	9.3588	0.50363	0.05463	9.2502	9.4675	8.5	10
Autonomy teachers	U1	161	8.1863	1.27208	0.10025	7.9883	8.3843	6	10
	C1	85	8.5765	1.00447	0.10895	8.3598	8.7931	6	10
Impact teachers	U1	161	8.9814	0.85675	0.06752	8.848	9.1147	6	10
	C1	85	9.1059	0.68615	0.07442	8.9579	9.2539	7.5	10
teacher LPI	U1	164	14.396	1.20466	0.09407	14.21	14.582	0	15
	C1	76	14.263	0.6175	0.07083	14.122	14.404	11	15
teacher SPES	U1	164	8.6416	0.95699	0.07473	8.494	8.7891	0	10
	C1	76	8.7199	0.65015	0.07458	8.5714	8.8685	6.93	10

Table 11. Test of normality for teacher LPI and SPES subscales. In this table, each of the LPI subscale categories is compared to a normal distribution using the one-sample Kolmogorov-Smirnov test. All of the subscales show significant ($p<0.05$) deviation from the normal distribution indicating that parametric tests which assume a normal distribution (t-tests, ANOVA, etc.) are not advised that instead non-parametric tests (Mann-Whitney, Kruskal-Wallis, etc.) be used.

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Model the way teachers is normal with mean 14.36 and standard deviation 0.551.	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
2	The distribution of Challenge the process teachers is normal with mean 14.20 and standard deviation 0.681.	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
3	The distribution of Enable others to act teachers is normal with mean 14.60 and standard deviation 0.499	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
4	The distribution of Encourage the heart teachers is normal with mean 14.45 and standard deviation 0.972	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
5	The distribution of Satisfaction teachers is normal with mean 8.52 and standard deviation 1.127.	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
6	The distribution of Decisionmaking teachers is normal with mean 7.92 and standard deviation 1.052.	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
7	The distribution of Professional growth teachers is normal with mean 8.93 and standard deviation 0.887.	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
8	The distribution of Status teachers is normal with mean 9.32 and standard deviation 0.757.	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
9	The distribution of Self efficacy teachers is normal with mean 9.27 and standard deviation 0.843.	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
10	The distribution of Autonomy teachers is normal with mean 8.35 and standard deviation 1.266.	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
11	The distribution of Impact teachers is normal with mean 8.96 and standard deviation 0.938.	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
12	The distribution of teacherLPI is normal with mean 14.37 and standard deviation 0.893.	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.
13	The distribution of teacherSPES is normal with mean 8.67 and standard deviation 0.779.	One-Sample Kolmogorov-Smirnov Test	.000[1]	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[1] Lilliefors Corrected

Table 12. Mann-Whitney ranks across districts U1 and C1 for teacher LPI and SPES subscale categories. Ranks across the 247 teachers from the two districts U1/C1) were assigned for all teachers' responses for each LPI, satisfaction, and SPES subscale category, this table shows both the average (mean) rank for each district as well as the sum of ranks for each district.

	District	N	Mean Rank	Sum of Ranks
Model the way teachers	U1	161	124.3	20019
	C1	85	121.9	10362
	Total	246		
Challenge the process teachers	U1	161	132.1	21263
	C1	85	107.3	9118
	Total	246		
Enable others to act teachers	U1	161	126	20290
	C1	85	118.7	10091
	Total	246		
Encourage the heart teachers	U1	161	126.5	20359
	C1	85	117.9	10023
	Total	246		
Satisfaction teachers	U1	161	106.3	17121
	C1	85	156	13261
	Total	246		
Decision-making teachers	U1	161	116.6	18772
	C1	85	136.6	11610
	Total	246		
Professional growth teachers	U1	161	111	17869
	C1	85	147.2	12512
	Total	246		
Status teachers	U1	161	118	18997
	C1	85	133.9	11385
	Total	246		
Self-efficacy teachers	U1	161	125.8	20252
	C1	85	119.2	10130
	Total	246		
Autonomy teachers	U1	161	116.6	18767
	C1	85	136.6	11614
	Total	246		
Impact teachers	U1	161	121.1	19489
	C1	85	128.2	10893
	Total	246		
Teacher LPI avg	U1	162	128.8	20865
	C1	85	114.9	9763.5
	Total	247		
Teacher SPES avg	U1	162	114.2	18500
	C1	85	142.7	12128
	Total	247		

Table 13. Test statistics comparing LPI and SPES subscale categories between district U1 and C1 teachers. This table shows the test statistics for the Mann-Whitney test performed across districts U1 and C1 for each sub-scale category and overall LPI and SPES scores. Statistical significance is represented by the "Asymp. Sig. (2-tailed)" row and statistically significant results are those with values of $p<0.05$. District U1 demonstrate significantly ($p=0.008$) higher scores for the LPI sub-score "challenge the process", but none of the other LPI categories or the overall LPI demonstrate significant differences. District C1 had significantly higher satisfaction scores ($p<0.001$) than U1. Furthermore, C1 also demonstrated significantly ($p=0.003$) higher SPES scores overall and higher SPES sub-scores in the areas of decision-making ($p=0.035$) professional growth ($p<0.001$), and autonomy ($p=0.034$).

Test Statistics[a]

	Model the way	Challenge the process	Enable others to act	Encourage the heart	Satisfaction	Decision-making	Professional growth	Status	Self-efficacy	Autonomy	Impact	LPI avg	SPES avg
Mann-Whitney U	6707	5463	6436	6367.5	4079.5	5730.5	4828	5955.5	6474.5	5726.0	6447.50	6108.50	5297.00
Wilcoxon W	10362	9118	10091	10022.5	17120.5	18771.5	17869	18996.5	10129.5	18767.0	19488.50	9763.50	18500.00
Z	-.263	-2.643	-.807	-.939	-5.232	-2.103	-3.883	-1.747	-.727	-2.124	-.786	-1.459	-2.979
Asymp. Sig. (2-tailed)	.792	.008	.419	.348	.000	.035	.000	.081	.467	.034	.432	.144	.003

a. Grouping Variable: District

Findings For Question 2b: Differences Between Teacher Leadership, Satisfaction, And Empowerment Across All Participating U.S. And Canadian Districts

In order to determine whether there were any differences in teacher leadership, satisfaction, and empowerment between the two countries, I first assessed the means and standard deviations for the relevant 4 LPI subscale sections, satisfaction, and the six SPES subscale sections (results are shown in the descriptive statistics, table 14). A Kolmogorov-Smirnov test was then used to assess the normality of the distributions for each category. All categories demonstrated significant ($p<0.001$) deviation from a normal distribution, and therefore non-parametric tests needed to be used to assess differences between districts. In order to assess differences in LPI/satisfaction/SPES scores, I performed a Mann-Whitney U test comparing the ranks of teachers' section scores between the two countries (213 teachers from the U.S. and 163 from Canada). The ranks table (Table 15) shows the mean rank and sum of ranks for each sub-scale category. To determine if the differences in ranks observed were significant, the 2-tailed p-values calculated by the Mann-Whitney test were calculated and are listed in the test statistics table (Table 16). No significant differences were in the 4 LPI subscale categories or in overall LPI scores. However, satisfaction scores were significantly higher amongst Canadian teachers compared to U.S. teachers (mean rank of 223.66 vs 161.59, $p<0.001$). The overall SPES score was significantly higher in Canada than in the U.S. (mean rank 214.53 vs 169.56, $p<0.001$) and the Canadian teachers also showed significantly higher scores in 3 of the 6 SPES sub-scale categories, "professional growth" ($p=0.001$), "autonomy" ($p<0.001$) and "decision-making" ($p=0.028$). Given these differences, *there are statistically significant differences between teachers' satisfaction and SPES empowerment categories between the two districts, but not leadership practices as measured by LPI*. These results indicate that I must reject the null

hypothesis that there are no significant differences between satisfaction, and SPES scores between the U.S. and Canadian teachers.

To assess if there were differences in any of the 30 individual questions, I also performed Mann-Whitney tests for each of the 30 individual items administered to the teachers as in 2a. 11 of the 30 questions showed significant differences with U.S. teachers having higher scores on 1 LPI question ("I challenge students to try new and innovative ways of learning") and Canadian teachers having higher scores on 1 ("I treat students with dignity and respect"), and Canadians having higher scores on all 4 satisfaction questions, and 4 SPES questions ("Principals and other teachers solicit my advice", "I work at a school where students come first", "I am able to teach as I choose", and "I have the freedom to make decisions on what is taught" (see supplemental table 4).

Table 14. Descriptive statistics for teacher SPES subscales across 8 participant districts. In this table mean, standard deviation, and range averages are shown for each of the LPI, satisfaction, and SPES subscale categories in each of the 8 participating US and Canadian districts

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
Model the way teachers	U1	161	14.449	0.46884	0.03695	14.3763	14.522	13	15
	U2	46	14.203	0.53277	0.07855	14.0447	14.361	12	15
	U3	6	13.333	1.26491	0.5164	12.0059	14.661	11	14.33
	U4	0	.	.	.	.	.	.	.
	C1	85	14.416	0.56789	0.0616	14.2932	14.538	11	15
	C2	47	14.355	0.38325	0.0559	14.2421	14.467	13.67	15
	C3	24	14.208	0.52762	0.1077	13.9855	14.431	13.33	15
	C4	7	14	1.10554	0.41786	12.9775	15.023	11.67	15
	Total	376	14.358	0.55134	0.02843	14.3022	14.414	11	15
Challenge the process teachers	U1	161	14.364	0.59489	0.04688	14.2718	14.457	12.67	15
	U2	46	14.007	0.61459	0.09062	13.8247	14.19	12	15
	U3	6	13.389	1.20031	0.49002	12.1292	14.649	11	14.33
	U4	0	.	.	.	.	.	.	.
	C1	85	14.118	0.74483	0.08079	13.957	14.278	11	15
	C2	47	14.149	0.63249	0.09226	13.9632	14.335	12.67	15
	C3	24	14.125	0.69374	0.14161	13.8321	14.418	12.67	15
	C4	7	14.095	0.99469	0.37596	13.1753	15.015	12	15
	Total	376	14.202	0.68068	0.0351	14.1331	14.271	11	15
Enable others to act teachers	U1	161	14.64	0.43297	0.03412	14.5724	14.707	13	15
	U2	46	14.457	0.55936	0.08247	14.2904	14.623	12	15
	U3	6	13.889	1.45551	0.59421	12.3614	15.416	11	15
	U4	0	.	.	.	.	.	.	.
	C1	85	14.612	0.51137	0.05547	14.5015	14.722	11	15
	C2	47	14.688	0.36385	0.05307	14.5811	14.795	13.67	15
	C3	24	14.722	0.32103	0.06553	14.5867	14.858	14	15
	C4	7	14.333	0.66667	0.25198	13.7168	14.95	13	15

	Total	376	14.605	0.49858	0.02571	14.5541	14.655	11	15
Encourage the heart teachers	U1	161	14.559	0.53308	0.04201	14.476	14.642	13	15
	U2	46	14.536	0.59826	0.08821	14.3586	14.714	12	15
	U3	6	13.722	1.46692	0.59887	12.1828	15.262	11	15
	U4	0	.	.	.	.	.	.	.
	C1	85	14.443	0.69463	0.07534	14.2933	14.593	11	15
	C2	47	14.518	0.55954	0.08162	14.3534	14.682	13	15
	C3	24	14.292	0.6469	0.13205	14.0185	14.565	12.67	15
	C4	7	12.429	5.51669	2.08511	7.3265	17.531	0	15
	Total	376	14.455	0.97207	0.05013	14.3562	14.553	0	15
Satisfaction teachers	U1	161	8.3789	1.04474	0.08234	8.2163	8.5415	6	10
	U2	46	7.9022	1.10865	0.16346	7.5729	8.2314	6	9.75
	U3	6	8.4167	1.20069	0.49018	7.1566	9.6767	6.5	10
	U4	0	.	.	.	.	.	.	.
	C1	85	9.0824	0.78337	0.08497	8.9134	9.2513	6.75	10
	C2	47	8.7766	0.86482	0.12615	8.5227	9.0305	6.5	10
	C3	24	8.4375	1.09904	0.22434	7.9734	8.9016	6.25	10
	C4	7	7.6429	3.38458	1.27925	4.5126	10.773	0	9.25
	Total	376	8.5199	1.12662	0.0581	8.4057	8.6342	0	10
Decision-making teachers	U1	161	7.9752	0.97836	0.07711	7.8229	8.1274	6	10
	U2	46	7.3098	0.96204	0.14184	7.0241	7.5955	6	9
	U3	6	8.2917	0.91401	0.37314	7.3325	9.2509	6.75	9.25
	U4	0	.	.	.	.	.	.	.
	C1	85	8.2382	0.82907	0.08993	8.0594	8.4171	6.5	10
	C2	47	7.867	1.01597	0.1482	7.5687	8.1653	6	10
	C3	24	7.8542	0.81065	0.16547	7.5119	8.1965	6	9.25
	C4	7	7	3.22426	1.21865	4.0181	9.9819	0	9
	Total	376	7.9189	1.05186	0.05425	7.8122	8.0255	0	10
Professional growth teachers	U1	161	8.8851	0.907	0.07148	8.7439	9.0263	6	10
	U2	46	8.5543	0.92032	0.13569	8.281	8.8276	6	10
	U3	6	8.75	0.93541	0.38188	7.7683	9.7317	7.5	10
	U4	0	.	.	.	.	.	.	.
	C1	85	9.3176	0.69799	0.07571	9.1671	9.4682	7	10
	C2	47	8.8723	0.82402	0.12019	8.6304	9.1143	6	10

	C3	24	8.7083	1.02062	0.20833	8.2774	9.1393	6	10
	C4	7	9.0714	0.7868	0.29738	8.3438	9.7991	8	10
	Total	376	8.9309	0.88725	0.04576	8.8409	9.0208	6	10
Status teachers	U1	161	9.3509	0.63454	0.05001	9.2522	9.4497	6	10
	U2	46	9.2391	0.63016	0.09291	9.052	9.4263	7.5	10
	U3	6	9.1667	0.68313	0.27889	8.4498	9.8836	8	10
	U4	0	.	.	.	.	.	.	.
	C1	85	9.5	0.49401	0.05358	9.3934	9.6066	7.5	10
	C2	47	9.234	0.46451	0.06776	9.0977	9.3704	8	10
	C3	24	9.2917	0.58823	0.12007	9.0433	9.5401	8	10
	C4	7	8	3.55903	1.34519	4.7084	11.292	0	10
	Total	376	9.3245	0.75704	0.03904	9.2477	9.4012	0	10
Self-efficacy teachers	U1	161	9.323	0.77944	0.06143	9.2017	9.4443	6	10
	U2	46	8.9457	0.73203	0.10793	8.7283	9.163	7	10
	U3	6	8.9167	0.58452	0.23863	8.3032	9.5301	8	9.5
	U4	0	.	.	.	.	.	.	.
	C1	85	9.3588	0.50363	0.05463	9.2502	9.4675	8.5	10
	C2	47	9.4149	0.53486	0.07802	9.2579	9.5719	8.5	10
	C3	24	9.3542	0.75871	0.15487	9.0338	9.6745	7	10
	C4	7	8.2857	3.67261	1.38812	4.8891	11.682	0	10
	Total	376	9.2726	0.84271	0.04346	9.1872	9.3581	0	10
Autonomy teachers	U1	161	8.1863	1.27208	0.10025	7.9883	8.3843	6	10
	U2	46	7.6848	1.26209	0.18609	7.31	8.0596	6	10
	U3	6	9.0833	0.91742	0.37454	8.1206	10.046	8	10
	U4	0	.	.	.	.	.	.	.
	C1	85	8.5765	1.00447	0.10895	8.3598	8.7931	6	10
	C2	47	8.9043	0.86375	0.12599	8.6506	9.1579	6	10
	C3	24	8.75	0.97802	0.19964	8.337	9.163	6	10
	C4	7	7.8571	3.4966	1.32159	4.6233	11.091	0	10
	Total	376	8.3471	1.26645	0.06531	8.2187	8.4755	0	10
Impact teachers	U1	161	8.9814	0.85675	0.06752	8.848	9.1147	6	10
	U2	46	8.7283	0.77249	0.1139	8.4989	8.9577	6.5	10
	U3	6	9.0833	0.80104	0.32702	8.2427	9.924	8	10
	U4	0	.	.	.	.	.	.	.
	C1	85	9.1059	0.68615	0.07442	8.9579	9.2539	7.5	10
	C2	47	9.0745	0.81413	0.11875	8.8354	9.3135	6	10
	C3	24	8.9375	0.98149	0.20035	8.5231	9.3519	6	10

	C4	7	7.6429	3.4966	1.32159	4.409	10.877	0	10
	Total	376	8.9641	0.93846	0.0484	8.8689	9.0593	0	10
teacherLPI	U1	164	14.396	1.20466	0.09407	14.2101	14.582	0	15
	U2	43	14.432	0.38329	0.05845	14.3142	14.55	13.58	15
	U3	21	14.448	0.36844	0.0804	14.2807	14.616	13.75	15
	U4	3	14.5	0.46398	0.26788	13.3474	15.653	14.08	15
	C1	76	14.263	0.6175	0.07083	14.1221	14.404	11	15
	C2	42	14.484	0.31356	0.04838	14.3864	14.582	13.67	15
	C3	20	14.308	0.84297	0.18849	13.9138	14.703	11	15
	C4	8	14.021	0.84721	0.29953	13.3126	14.729	12	14.75
	Total	377	14.374	0.89314	0.046	14.2838	14.465	0	15
teacherSPES	U1	164	8.6416	0.95699	0.07473	8.494	8.7891	0	10
	U2	43	8.3837	0.65633	0.10009	8.1817	8.5857	7.07	9.71
	U3	21	8.4898	0.48984	0.10689	8.2668	8.7128	7.71	9.43
	U4	3	8.4524	0.75705	0.43708	6.5718	10.333	7.64	9.14
	C1	76	8.7199	0.65015	0.07458	8.5714	8.8685	6.93	10
	C2	42	8.8776	0.51102	0.07885	8.7183	9.0368	7.64	9.93
	C3	20	8.9997	0.39876	0.08917	8.8131	9.1864	8.07	9.71
	C4	8	8.7589	0.25308	0.08948	8.5473	8.9705	8.43	9.07
	Total	377	8.6658	0.77855	0.0401	8.5869	8.7446	0	10

Table 15. Cross-national teacher ranks in LPI and SPES sub scores. Ranks across the 377 teachers from the two countries (four districts/country) were assigned for teachers' responses for each LPI, satisfaction, and SPES subscale category, this table shows both the average (mean) rank for teachers from each country as well as the sum of ranks for each country listed by sub-scale item.

	Country	N	Mean Rank	Sum of Ranks
Model the way teachers	US	213	190.8	40635
	Canada	163	185.5	30242
	Total	376		
Challenge the process teachers	US	213	197.5	42061
	Canada	163	176.8	28816
	Total	376		
Enable others to act teachers	US	213	185.1	39416
	Canada	163	193	31461
	Total	376		
Encourage the heart teachers	US	213	194.9	41519
	Canada	163	180.1	29357
	Total	376		
Satisfaction teachers	US	213	161.6	34420
	Canada	163	223.7	36457
	Total	376		
Decision-making teachers	US	213	177.7	37857
	Canada	163	202.6	33020
	Total	376		
Professional growth teachers	US	213	172.8	36815
	Canada	163	209	34062
	Total	376		
Status teachers	US	213	186.3	39690
	Canada	163	191.3	31187
	Total	376		
Self-efficacy teachers	US	213	182.8	38933
	Canada	163	196	31944
	Total	376		

Autonomy teachers	US	213	167	35576
	Canada	163	216.6	35301
	Total	376		
Impact teachers	US	213	181.3	38623
	Canada	163	197.9	32254
	Total	376		
Teacher LPI avg	US	214	194.6	41640
	Canada	163	181.7	29614
	Total	377		
Teacher SPES avg	US	214	169.6	36285
	Canada	163	214.5	34968
	Total	377		

Table 16. Test statistics for cross-national comparison of teacher LPI and SPES sub-scale categories. This table shows the test statistics for the Mann-Whitney test performed across US and Canadian teachers for each LPI, satisfaction, and SPES sub-scale category and overall LPI and SPES scores. Statistical significance is represented by the "Asymp. Sig. (2-tailed)" row and statistically significant results are those with values of $p<0.05$. No significant differences in LPI were observed. Canadian teachers had significantly higher satisfaction scores ($p<0.001$) than US teachers. Furthermore, Canadian teachers also demonstrated significantly ($p<0.001$) higher SPES scores overall and higher SPES sub-scores in the areas of decision-making ($p=0.028$) professional growth ($p=0.001$), and autonomy ($p<0.001$).

	Model the way	Challenge the process	Enable others to act	Encourage the heart
Mann-Whitney U	16875.5	15449.5	16624.5	15991
Wilcoxon W	30241.5	28815.5	39415.5	29357
Z	-0.475	-1.855	-0.739	-1.368
Asymp. Sig. (2-tailed)	0.635	0.064	0.46	0.171

	Satisfaction	Decision-making	Professional Growth	Status	Self-Efficacy	Autonomy	Impact	LPI avg	SPES avg
Mann-Whitney U	11628.5	15065.5	14023.5	16898.5	16141.5	12784.5	15831.5	16247.5	13280
Wilcoxon W	34419.5	37856.5	36814.5	39689.5	38932.5	35575.5	38622.5	29613.5	36285
Z	-5.512	-2.204	-3.259	-0.46	-1.216	-4.427	-1.55	-1.141	-3.972
Asymp. Sig. (2-tailed)	0.000	0.028	0.001	0.645	0.224	0.000	0.121	0.254	0.000

3. What relationships exist between administrative and teacher leadership practices?

Findings for question 3: Correlation between administrative and teacher LPI scores

To investigate if there was a statistically significant association between administrator LPI sub-scores and teacher LPI sub-scores from teachers and administrators at the same school, a correlation was computed. Both of these distributions are non-normal (see K-S test calculated in questions 1 and 2) which violates the assumption of normality and therefore I calculated the Spearman rho statistic to estimate correlation. Because not all schools had teacher and administrator participant, data were limited to the 26 schools in which at least one teacher and one administrator completed the survey (16 of these schools had >8 teachers and >3 administrators respond). Each individual teacher's LPI subscale categories was correlated with average administrator LPI sub scores for their school, and this was performed (1) among all schools (2) among U.S. schools and (3) among Canadian schools. Among all schools, significant negative correlations were seen between teacher's "challenge the process" with administrator's "challenge the process" and "inspire a shared vision" sub-scores (Table 17). Conversely, teacher's "encourage the heart" sub-score positively and significantly correlated with their school administrator's "model the way" sub-score. According to Cohen's (1988) guidelines the effect size of each of these correlations is small.

When breaking the correlations down by country, the U.S. teachers' "encourage the heart" scores were significantly positively correlated with 4/5 administrator LPI sub-scores including "model the way", "challenge the process", and "enable others to act" with small to medium effect sizes (Table 18). The Canadian administrative LPI sub-scores did not correlate with any of the teacher LPI sub-scores (Table 19).

Table 17. Correlations between teacher LPI scores and administrator LPI scores. Schools in which at least 1 teacher and 1 administrator completed the survey (26 total) were included. Significant negative correlations were seen between administrator "challenge the process" and "inspire a shared vision" scores with teacher "challenge the process" scores while there was a significant positive correlation between administrator's "model the way" scores with teachers' "encourage the heart" sub-score.

Spearman's Rho			Model the way A	Challenge the process A	Enable others to act A	Encourage the heart A	Inspire a shared vision A
	Model the way T	Correlation Coefficient	-0.054	-0.058	-0.024	-0.059	-0.074
		Sig. (2-tailed)	0.448	0.415	0.737	0.407	0.303
		N	197	197	197	197	197
	Challenge the process T	Correlation Coefficient	-0.048	-.142*	-0.048	-0.099	-.163*
		Sig. (2-tailed)	0.5	0.046	0.507	0.165	0.022
		N	197	197	197	197	197
	Enable others to act T	Correlation Coefficient	0.056	0.017	0.058	0.051	-0.017
		Sig. (2-tailed)	0.436	0.811	0.421	0.478	0.809
		N	197	197	197	197	197
	Encourage the heart T	Correlation Coefficient	.146*	0.034	0.114	0.086	-0.002
		Sig. (2-tailed)	0.04	0.634	0.111	0.23	0.978
		N	197	197	197	197	197

Table 18. Correlations between teacher LPI scores and administrator LPI scores from their school across US schools in which at least 1 teacher and 1 administrator completed the survey. Significant positive correlations were seen between four administrator LPI sub-scores with the teachers' responses to the "encourage the heart" sub-scale.

			Model the way A	Challenge the process A	Enable others to act A	Encourage the heart A	Inspire a shared vision A
Spearman's rho	Model the way T	Correlation Coefficient	-0.005	-0.02	0.064	0.033	0.177
U.S.		Sig. (2-tailed)	0.961	0.854	0.56	0.761	0.103
		N	86	86	86	86	86
	Challenge the process T	Correlation Coefficient	0.051	0.032	0.092	0.029	0.124
		Sig. (2-tailed)	0.638	0.769	0.4	0.791	0.254
		N	86	86	86	86	86
	Enable others to act T	Correlation Coefficient	0.089	0.15	0.073	0.084	0.053
		Sig. (2-tailed)	0.418	0.169	0.505	0.442	0.63
		N	86	86	86	86	86
	Encourage the heart T	Correlation Coefficient	.298**	.256*	.290**	.277**	0.162
		Sig. (2-tailed)	0.005	0.017	0.007	0.01	0.136
		N	86	86	86	86	86

Table 19. Correlations between teacher LPI scores and administrator LPI scores across all Canadian schools in which at least 1 teacher and 1 administrator completed the survey. No significant correlations were observed.

			Model the way A	Challenge the process A	Enable others to act A	Encourage the heart A	Inspire a shared vision A
Spearman's rho	Model the way T	Correlation Coefficient	-0.037	-0.004	-0.036	0.017	-0.151
Canada		Sig. (2-tailed)	0.701	0.967	0.706	0.86	0.113
		N	111	111	111	111	111
	Challenge the process T	Correlation Coefficient	0.007	-0.057	0.007	0.06	-0.091
		Sig. (2-tailed)	0.941	0.55	0.941	0.531	0.344
		N	111	111	111	111	111
	Enable others to act T	Correlation Coefficient	0.044	-0.056	0.055	0.056	-0.027
		Sig. (2-tailed)	0.65	0.558	0.57	0.557	0.776
		N	111	111	111	111	111
	Encourage the heart T	Correlation Coefficient	0.109	0.003	0.087	0.118	-0.038
		Sig. (2-tailed)	0.254	0.972	0.363	0.216	0.692
		N	111	111	111	111	111

4. What relationships exist between administrative leadership practices and teacher empowerment?

Findings for question 4: Correlation between administrative and teacher SPES scores

To investigate if there was a statistically significant association between administrator LPI sub-scores and teacher SPES sub-scores from teachers and administrators at the same school, a correlation was computed. Both of these distributions are non-normal (see K-S test calculated in questions 1 and 2) which violates the assumption of normality and therefore I calculated the Spearman rho statistic to estimate correlation. Because not all schools had teachers and administrator participants, data were limited to 26 schools in which at least one teacher and one administrator completed the survey. While this represents only a subset of schools, it provides useful data for all schools in which administrators' LPI could influence teacher responses. Each individual teacher's SPES subscale categories was correlated with average administrator LPI sub scores for their school, and this was performed (1) among all schools (2) among U.S. schools and (3) among Canadian schools. Among all schools, significant positive correlations were seen between teacher's "autonomy" and administrator's "challenge the process", "encourage the heart" and "inspire a shared vision" sub-scores. Conversely, teacher's "status" SPES sub-score negatively and significantly correlated with their school administrator's "model the way" sub-score with a small effect size (Table 20).

When breaking the correlations down by country, the U.S. teachers' "self-efficacy" scores were significantly positively correlated with administrator "inspire a shared vision" LPI sub-score with a medium effect size (Table 21). Like the overall analysis, significant positive correlations were seen between U.S. teacher's "autonomy" and administrator's "challenge the process", "encourage the heart" and "inspire a shared vision" sub-scores, these showed small to

medium effect sizes. Canadian teacher's "status" SPES sub-score significantly and negatively correlated with "model the way", "enable others to act" and "inspire a shared vision" administrative LPI sub-scores while teacher decision-making negatively correlated with the administrator's "inspire a shared vision" sub-score (Table 22).

Table 20. Correlations between teacher SPES scores and administrator LPI scores. School in which at least 1 teacher and 1 administrator completed the survey (26 total) were included. Significant negative correlations were seen between administrator "model the way" scores with teacher "status" scores while there was a significant positive correlation between administrator's "inspire a shared vision", "encourage the heart", and "challenge the process" scores with teachers' "autonomy" sub-score.

Spearman's rho		Model the way A	Challenge the process A	Enable others to act A	Encourage the heart A	Inspire a shared vision A
Decision-making	Correlation Coefficient	-0.071	0.061	-0.014	0.023	0.004
	Sig. (2-tailed)	0.318	0.397	0.849	0.753	0.952
	N	197	197	197	197	197
Professional growth	Correlation Coefficient	-0.06	0.1	0	-0.013	0.066
	Sig. (2-tailed)	0.4	0.162	1	0.857	0.354
	N	197	197	197	197	197
Status	Correlation Coefficient	-.184**	-0.046	-0.136	-0.115	-0.041
	Sig. (2-tailed)	0.01	0.524	0.057	0.109	0.563
	N	197	197	197	197	197
Self efficacy	Correlation Coefficient	-0.027	0.076	-0.015	0.031	0.117

	Sig. (2-tailed)	0.702	0.292	0.832	0.666	0.102
	N	197	197	197	197	197
Autonomy	Correlation Coefficient	0.132	.244**	0.116	.202**	.263**
	Sig. (2-tailed)	0.064	0.001	0.103	0.004	0.000
	N	197	197	197	197	197
Impact	Correlation Coefficient	-0.073	0.028	-0.045	-0.02	0.012
	Sig. (2-tailed)	0.311	0.696	0.527	0.779	0.863
	N	197	197	197	197	197

Table 21. Correlations between teacher SPES scores and administrator LPI scores in US schools.

Spearman's rho - US		Model the way A	Challenge the process A	Enable others to act A	Encourage the heart A	Inspire a shared vision A
Decision-making	Correlation Coefficient	-0.009	0.173	0.066	0.165	0.211
	Sig. (2-tailed)	0.935	0.111	0.544	0.129	0.051
	N	86	86	86	86	86
Professional growth	Correlation Coefficient	-0.054	0.015	0.019	-0.017	0.027
	Sig. (2-tailed)	0.619	0.888	0.864	0.88	0.807
	N	86	86	86	86	86
Status	Correlation Coefficient	-0.205	-0.097	-0.154	-0.148	0.067
	Sig. (2-tailed)	0.058	0.373	0.156	0.174	0.542
	N	86	86	86	86	86
Self efficacy	Correlation Coefficient	-0.034	0.096	0.024	0.066	.330**
	Sig. (2-tailed)	0.755	0.379	0.827	0.545	0.002
	N	86	86	86	86	86
Autonomy	Correlation Coefficient	0.119	.282**	0.082	.232*	.288**
	Sig. (2-tailed)	0.275	0.008	0.452	0.031	0.007
	N	86	86	86	86	86
Impact	Correlation Coefficient	-0.057	0.089	-0.011	0.038	0.138
	Sig. (2-tailed)	0.603	0.417	0.921	0.726	0.204
	N	86	86	86	86	86

Table 22. Correlations between teacher SPES scores and administrator LPI scores in Canadian schools.

Spearman's rho Canada		Model the way A	Challenge the process A	Enable others to act A	Encourage the heart A	Inspire a shared vision A
Decision-making	Correlation Coefficient	-.159	-.094	-.143	-.111	-.237*
	Sig. (2-tailed)	.095	.326	.135	.246	.012
	N	111	111	111	111	111
Professional growth	Correlation Coefficient	-.177	-.023	-.168	-.168	-.138
	Sig. (2-tailed)	.063	.812	.077	.079	.150
	N	111	111	111	111	111
Status	Correlation Coefficient	-.213*	-.126	-.202*	-.166	-.222*
	Sig. (2-tailed)	.025	.186	.034	.081	.019
	N	111	111	111	111	111
Self efficacy	Correlation Coefficient	-.017	.063	-.036	.032	.058
	Sig. (2-tailed)	.858	.508	.706	.741	.545
	N	111	111	111	111	111
Autonomy	Correlation Coefficient	.060	.075	.037	.048	.176
	Sig. (2-tailed)	.533	.434	.697	.618	.065
	N	111	111	111	111	111
Impact	Correlation Coefficient	-.079	-.051	-.086	-.061	-.075
	Sig. (2-tailed)	.410	.599	.369	.523	.432
	N	111	111	111	111	111

E. DISCUSSION AND INTERPRETATION OF STATISTICAL RESULTS

Study Orientation

The current study builds heavily on and is motivated by the notion that that U.S. schools lag behind other similar industrialized nations and can learn from other high-achieving nations as advanced by authors such as Darling-Hammond (2011) and Baker & LeTendre (2005). By comparing urban schools in four districts from the U.S. and four similar districts from Canada, the nation whom the closest comparison can be made (Hargreaves & Shirley, 2012), this work adds to the field of comparative education (Bray & Adamson, 2007). *Instead of, however, focusing on comparative education in terms of outcomes such as test scores, this work focuses on the importance of teacher and administrator leadership which scholars such as Dong and Cravens (2011) note is conspicuously absent from most international educational scholarship*. As my study joins concepts from the fields of international comparative education and educational leadership, I have used transformational leadership theory (Burns, 1978; Bass, 1985) as a theoretical framework to assess differences between teachers and administrators at urban schools in the United States and Canada. To do this, I have utilized and interwoven the work of Kouzes and Posner (1995) and Short and Rinehart (1992) to define and assess important concepts which underlie leadership practices and teacher empowerment. The present study is thus able to assess significant domains such as leadership behaviors amongst principals and teachers (as measured by the LPI), teacher empowerment (as measured by the SPES), and their inter-relationships in the international context.

In addition, while an important body of my review of literature focused on international tests, I am not interested in comparing schools from different nations on the basis of test scores because these score are unnecessarily deterministic and do not do a good job assessing the true essence of schooling (Bracey, 2009; Kell & Kell, 2010). Instead, my study examines leadership as a frequently under-studied aspect of schooling which may have important cross-national implications. Something which is not well reflected by tests, leadership behaviors impact schools' mission, vision, culture, and success (Dong & Cravens, 2011). Whether and how leadership amongst teachers and administrators differs in the cross-national context, therefore, is an important topic to study.

Response rates, tiered study design

A total of 460 individuals who fully completed my survey, of these 257 were from the US and 203 were Canadian. The largest districts from each country were U1 with 181 responses and C1 with 105 responses. None of the individual districts had a survey response rate of >10% (range 1-8.5%) when compared to the population of all eligible teachers and administrators from each district. However, methodological issues beyond study control such as lack of availability and out-of-date of email addresses, lack of principal forwarding recruitment emails to staff, and district firewalls which blocked the survey were all factors in the low response rates seen. Given these issues, while the response rates are low compared to what can be expected with in-person and paper surveys (often 25 to 50 percent), but are consistent with previous studies using online surveys which are often twenty percent lower than paper surveys (Nulty, 2008). A 2 district cross-national comparison was my original study design, but due to need to apply to multiple districts in order to obtain adequate numbers of respondents and receiving responses from eight

participating districts, my study became (1) a primary study comparing two districts, one in the US and one in Canada and (2) a study comparing results between the four participant US districts with the four participating Canadian districts. The additional analysis is valuable not only in that it provides additional cross-district comparisons within each country, but it also increases the sample size and therefore improves statistical power of the study.

The reasons for non-response cited by teachers, administrators, and district personnel were largely things such as time, convenience, and other responsibilities. Because the survey did not ask for sensitive information or collect personally identifiable information, the non-response observed appears to be due to logistical factors and it is reasonable to assume that the non-response was not significantly influenced by the questions asked on the survey. In addition, there were a small number of individuals (n=45) who began the survey but did not complete it (incomplete responders). These individuals were excluded from the final analysis, and their differing stages of dropout, time spent on the survey, and small numbers (<10% of total participants) suggest that there was not one part of the survey which specifically led participants to give an incomplete response and/or drop out.

Districts and School Boards in study

Cross-National Comparison: Selection of Eight Participant Districts

A total of 51 US districts and Canadian School Boards were originally identified as potential subjects for my study on the basis that they had city populations of >300,000 and at least 30,000 students enrolled in their public schools. As the flow chart in Figure 1 shows (on p.5), a number of these schools were not selected because they were not accepting applications to

perform research from external applicants or had a process that would preclude their participation during the study's time frame. Of the thirty-two districts that remained potentially eligible, complete applications were submitted to each district. Applications were accepted at twenty-five districts, and after numerous rounds of modifications based on the individual review and clearance process, eight of these were approved and the study could be performed during the request timeframe. Applications to perform research were not accepted at twenty-six (76%) of these districts, but the reasons for district rejections were highly variable. Multiple districts stated that their teachers were already overburdened, some noted that the district was already performing other research about leadership (although not international) or using online survey instruments (although not LPI/SPES) and they viewed this research as similar, some commented that they did not see relevance of the study to their district priorities (because the results would not have been immediately measurable), some did not like being compared with another district, some did not think that selection of regular/college preparatory schools was acceptable (they would only accept a study of all schools), some required a paper version of the signed consent from a building Principal in order for a school to participate and afterwards, required that each participant did not remain anonymous and also had to sign a paper consent, some districts thought the number of schools requested for research was too many, and some were not comfortable with participants needing to disclose which school they worked in, and one district did not want to include vice principals as administrators because they only viewed principals as administer-leaders. Given the high variability of why districts chose not to participate, and the self-selection of participant districts, there was *no systematic bias* in district selection and the eight schools that were included in the study can be considered to be a random sample of the 64

(51 US / 13 Canadian) potential participant districts. Given this context, in using inferential statistics, the comparison of 4 U.S. districts with 4 Canadian districts can be considered a random sample of a cross-national comparison of the intended population. This concept of the eight districts sampled representing a random sample of the total population of eligible districts/school boards makes the non-response of the other districts/boards less problematic in that the results of the population (all 51 US districts and 13 Canadian boards) are inferred to be statistically represented by the eight sampled districts.

The eight individual districts that did participate each had (after rounds of study modification through their individual review processes) different requirements and these may have had some impact on study participation. Some districts allowed me to send emails to all administrators and teachers (using whatever publicly available email addresses that I could find) while others only allowed me to contact principals who were then asked to (but did not necessarily) contact teachers. Some districts required individual principals to approve of the study, and one school board did not allow respondents to indicate their school. While these differences meant that some districts had higher administrator responses and lower teacher responses or vice-versa, and they meant that the one district with no school level identifications (C4) had to be excluded from the analysis of questions 3 and 4, none of these factors should have impacted the way in which participants would have answered the survey questions and therefore should not compromise the integrity of the data.

One concern about cross-district and cross-national comparisons is the issue that districts were drawn from different states/provinces within the U.S. and Canada and that each of the states and provinces has different local political, economic, and educational contexts (Hargreaves

& Shirley, 2012). Moreover, even within individual states and provinces local districts and school boards have different policies, leadership initiatives, and contexts which make the process of cross-district comparison (whether it be 2 districts or 8 districts) inherently challenging. In Canada, much of the literature about successes in Canadian education including work from Hargreaves & Shirley (2012) focuses primarily on studies done within Ontario. In the U.S., while multiple districts have been widely studied, the political/educational, and economic differences across states and cities has been shown to be a significant factor in differences in educational outcomes (Dunne & Fee, 2008). On this basis, one could argue that comparisons should instead be framed as "a large district in Texas compared to a large district in Alberta", for example, or "four large U.S. districts drawn from New York, Texas, and California, and four large Canadian school boards from Ontario, Alberta, and Manitoba". Such state/province level differences, however, because they can only account for some aspects of education at a policy-level, would still fail to account for local contexts both of the individual districts as well as the schools sampled within the districts (Datnow, 2008).

Despite these inherent truths, there is an imperative to compare education in different locales (Kelleghan 2006). While there is not established literature specifically saying that, for example, the states/provinces or the specific school districts U1 and C1 are "equivalent" (and this would be an unreasonable standard), to address the local contexts, I have carefully selected schools on the basis of size, school number, geography, and economy to minimize local differences from being the main driver of effects seen. Furthermore, statistical procedures including correction for state level, meta-analysis, and within-country analysis can mitigate such

effects, but without a more in-depth analysis of local contexts, it is hard to entirely account for such differences.

While "cross-national" (across the four U.S. districts and four Canadian districts) results were analyzed, it is important to remember that districts within the same country have some inherent differences and that grouping districts together by country, while adding statistical power, may introduce heterogeneity as an additional confounding factor in the cross-national analysis. Meta-analysis would be one way to overcome this limitation, but given the small response sample sizes from districts U3, U4, C3, and C4, these districts would likely have large confidence intervals and need to be excluded from such an analysis. Therefore, while the small sample sizes make these districts more prone to sampling bias when considered individually, their inclusion as part of the "cross-national" study allows me to analyze these responses (which are vital, especially to providing a sufficient number of administrators to compare) without substantially changing the overall meaning of the study question interpretation.

Another issue to be addressed is that the 8 districts (U1-U4 and C1-C4) are representative of 3 US states and 3 Canadian provinces; (U2 and U3 are from one state, C3 and C4 are from one province). Given that both the US and Canadian education systems are not centralized but rather give substantial local leeway to the administration of education policy to the state/provincial governments (Levin, 2011), differences at this level could influence educational outcomes including those assessed in this study. Since the state/province which participant responses come from cannot be truly "corrected for" statistically (other than including state/province a demographic covariate, which I performed, but results provided very little additional information because of study design), I instead assessed if there was substantial

variability of responses across districts within each country and whether this variability was significant. To do this, I split the data by country (i.e. analyzing US responses and then separately analyzing Canadian responses) and assessed for differences in each of the 9 LPI (5 administrator and 4 teacher sub-scales) and SPES (6 sub-scales) categories for teachers and administrators using an ANOVA test with a Tukey post-hoc analysis. While there were significant differences between respondents across districts, when data was re-coded by state/province, the post-hoc test did not reveal any statistically significant differences in any of the sub-scales for either country. This suggests that state/province-wide differences were not a primary driver of the associations observed. Based on this lack of inter-state/provincial difference and the notion that the districts represent a random sample of all eligible districts in the two countries, I contend that it is reasonable to combine the four districts/boards from each country instead of treating them separately.

Cross-District Comparison: U1 vs C1

The study was originally intended to be a comparison of two of the largest districts, one in the U.S. and one in Canada that were selected on the basis of size, education systems, diversity, and school type similarities to ensure an appropriate comparison was possible. Because one of the original districts did not approve the study and the other had a very low response rate, this original strategy needed to be modified and I applied to thirty-four districts/school boards to ensure that there would be adequate participation in the study. Within this setup, I identified at least one comparable US/Canadian district for each of the districts I applied to so that a cross-district comparison could be made on the basis of city size and eligible

participating schools. After the study closed, of the eight participant districts, only two districts had over 100 participants and both had 20 administrators. These two districts, U1 and C1, each were districts from cities of between one and two million residents and had similar numbers of schools and students in their districts. Both cities share a common history of industry centered on the same commodity. While district U1 did have almost twice as many teacher respondents, given the relatively large number of teachers in both districts, this discrepancy is not likely to change the meaning of the results as differences are assessed based on average scores and both are large enough to have a robust average score. One factor which may be worth considering in the cross-district comparison is that district U1 allowed me to contact teachers (using whatever publicly available emails that I could find) while district C1 required teachers to be contacted by their principal. This led me to have a large number of responses from a smaller number of schools in district C1 (five schools with 10 or more teacher responses) while district U1 had more schools with teacher responses (eighteen versus seven schools) but a lower number of responses per school. Review of literature regarding the state and province in which these districts are located revealed local differences including more focus on test-based accountability in U1's state and more focus on school choice in C1's province (Vazquez Hellig, Jez, & Reddickm, 2012; Milke, 2010).

Demographics of Study Participants

In order to determine whether there were any differences in gender, ethnicity, age, years of teaching experience, and level of education between the two districts, I first assessed the distribution of each using both graphical methods and using a Mann-Whitney U test (for all but

Ethnicity which had ordinal data and therefore needed to be compared using a Kruskal-Wallis test (supplemental tables 5-8). To determine if the differences in ranks observed were significant, the 2-tailed p-values determined by the Mann-Whitney test were compared and are listed in the test statistics table. Significant differences were seen in years of teaching experience ($p<0.001$) and level of education ($p<0.001$) but not for gender, age, or ethnicity. While there are apparent differences in ethnicity both across districts and cross-nationally, with the US districts generally having higher proportions of non-Caucasian participants, only the cross-national comparison of ethnicity (and not the cross-district analysis) demonstrated statistical significance.

The two demographic variables that showed the greatest difference both in the cross-district and the cross-national comparison were *years of teaching/administrative experience* and *level of education*. Interestingly, the Canadian participants had significantly more teaching experience (a substantially higher proportion of C1/Canadian participants had 10-20 years of experience while in the U1/US many more had only 0-10 years of experience). Whether the difference in experience has to do with burnout, satisfaction, retention, or other factors is unclear but implies Canadian educators are more likely to stay with their profession. The years of experience is in contrast to level of education where the average US participant completed a master's degree while the average Canadian participant had completed a bachelor's. This may represent differences in national/cultural norms or potentially other external factors (such as US districts paying more to teachers with a master's degree and no equivalent salary change in Canadian schools).

In addition to my primary analysis which looked only at whether there were differences in demographic variables, I performed additional analyses (supplemental tables 9-12) to assess

whether the demographic differences seen across districts U1/C1 and cross-nationally may account for some of the differences in leadership or empowerment seen in questions 1 and 2. To do this, I performed a post-hoc analysis that included the significant covariates (*years of experience* and *level of education* in the cross-district analysis and these plus *ethnicity* in the cross-national comparison) in addition to the primary LPI/SPES scores and grouping by district/country. Using this strategy, I found that demographic differences accounted for many of the differences seen in teacher empowerment sub-scales in the U1 vs C1 comparison but that the demographic differences did not affect the cross-national study results.

Non-Normality Of Results

Survey research which uses Likert scale questions frequently leads to non-normally distributed results (Sullivan & Artino, 2013), and therefore it is not unexpected that all survey questions asked demonstrated significant levels of skewness/kurtosis and showed significant deviation from normal upon testing with the Kolmogorov-Smirnov test. Given this type of distribution, non-parametric statistics were used in the interpretation of the results because these tests (Mann Whitney, Kruskal-Wallis, Spearman's rho, etc.) do not assume the presence of a normal distribution. When responses are clustered, parametric tests may potentially lead the mean to appear to be the neutral or middle response, and in doing so not fairly characterize the data or not represent a useful measure of the data's central tendency. To avoid this issue, the Spearman rho assessment and Mann-Whitney U test should be used for analysis of Likert scale questions instead of parametric tests that assume interval data.

Based on the central limit theorem (CLT), because I had a sufficiently large sample size ($n>30$) for all comparisons (other than administrators in U1 vs C1) it would still be possible to

perform analysis using parametric tests and they should be valid because for interval estimation, the standard error is computed from a sampling distribution of the mean and when the sample size approaches 30, the sampling distribution approaches normality (NIST/SEMATECH e-Handbook of Statistical Methods, 2012*).* I performed parametric analyses for questions 1 and 2 and did not see major differences between t-tests and Mann-Whitney tests which validates this concept and means the assumption of normality does not drastically change the study results. While the parametric tests have more statistical power than do non-parametric tests, to ensure that I meet all test assumptions, I only report non-parametric test results.

Question One and Two: Comparison of administrator leadership and teacher leadership, satisfaction, and empowerment across districts and across nations

Question 1 demonstrated that there were no significant differences between administrators' LPI scores across either districts U1 and C1 or across US and Canadian administrators. This lack of significance is interesting because it indicates that cross-district and cross-nationally, administrators self-report the same types of leadership behaviors. Given the well-described validity of the LPI (Kouzes & Posner, 2002), it seems reasonable to assume that the responses are valid and represent what administrators believe they do (demonstrated behavior).

Whether this report of (generally high) scores for each of the leadership sub-scores is reliable (i.e. whether the administrators actually do all of the things which they report that they do) cannot be determined in the present study. The administrator scores also do not differ substantially from previous implementations of the LPI amongst school principals (Richardson, Flanigan, Lane, & Keaster, 1992). Previous studies (Leech & Fulton, 2008; Shead, 2010;

Garity, 2011) have shown one way to do this is to have teachers rate their principal's leadership, but for this study I am more interested in whether the principal's self-reported leadership influenced teacher leadership and empowerment.

While some of the lack of statistical difference seen across districts and cross-nationally may reflect the smaller sample size of administrators and therefore reduced power, I did have >80% power to detect significant associations cross-nationally and still found no differences which leads me to conclude that there are no differences rather than that the study missed real differences.

Two of the thirty individual administrator LPI items ("I experiment and take risks even when there is a chance of failure" and "I praise teachers for a job well done") showed statistically significant differences with district C1 administrators reporting these practices far more frequently than administrators from district U1. These responses indicate that district C1 administrators are more willing both to experiment as well as to give teachers feedback even though they are not more likely to "challenge the process" or "encourage the heart" when all questions were considered. Whether these results stem more from administrator preference, district policies, culture or some other factor is hard to ascertain and may not have significant meaning given the overall lack of difference on the LPI instrument as a whole. In the cross-national comparison, 3 of the 30 items ("I follow through on the promises and commitments that I make", "I ask for feedback on how my actions affect teachers' performance", and "I praise teachers for a job well done") showed statistically significant differences. In this case, U.S. administrators reporting asking for feedback more frequently while Canadian administrators reported following through on commitments and praising teachers. The US administrators

asking for feedback may have to do with administrators' need to be held accountable for 'adequate yearly progress', to generate reports, and demonstrate the effect of their policies (Linn, 2011) while the Canadian administrators' demonstrated qualities more aligned with personal accountability and giving positive feedback which are not as easily quantifiable but are in line with literature suggesting recent shifts towards more transformational principal leadership in Canada (Fensterwald, 2013).

In question 2, multiple differences were seen between teacher leadership practices, satisfaction, and empowerment both across districts U1 and C1 and cross-nationally. Amongst LPI categories, U1 teachers had significantly higher scores than C1 teachers ($p=0.008$) in the "challenge the process" subscale. This is in contrast to the cross-national comparison where no statistically significant difference was seen ($p=0.06$). This discrepancy represents differences between districts within the countries. Because no demographic covariates explain this (see question 5), it seems likely that teachers from district U1 had a particularly high score in this category compared to all of the other 7 districts and there may be a factor not fully assessed in this study that leads the teachers in district U1 to be particularly likely to "challenge the process". According to Kouzes and Posner (1995), challenging the process entails searching for opportunities to change, grow, and improve in innovative ways. This means creating challenges for others as a leader seeks out meaningful challenges for him/herself (Kouzes & Posner, 1995). When there is rapport, trust, and respect between teachers and the principal, there is significantly higher likelihood of improved pedagogy and increased student achievement (Zimmerman & Deckert-Pelton, 2003).

The most significant difference that I consistently found was that Canadian teachers (including C1 teachers vs U1) were significantly more satisfied with their jobs, school leaders, and leadership team than were their American counterparts. This difference ($p<0.001$) was evident in all of the 8 districts although one American district (U2) had particularly low satisfaction scores. Across both districts U1 and C1 and the US and Canada, Canadian teachers showed statistically significant higher levels of satisfaction on all four individual satisfaction questions as well the aggregate category score. Because 'satisfaction' did not associate significantly with any of the administrator LPI scores (in contrast to previous work by Bogler, 2001), it seems likely that non-leadership related factors (including but not limited to school structure, teacher expectations, classroom load, district/national policies, etc.) likely have major contributions to higher satisfaction levels amongst district C1 and the Canadian teachers in this study. This is in line with scholarship that U.S. teachers who often work longer hours and work in isolation, are less satisfied with their jobs than teachers abroad (Sparks, 2014).

The overall SPES score was significantly higher in district C1 than in U1 and the teachers from this district also showed significantly higher scores in 3 of the 6 SPES sub-scale categories including professional growth, autonomy, and decision-making. This indicates that teachers from district C1 were significantly more likely than U1 teachers to feel that their schools provided them with opportunities to function independently, make decisions that impact their schools, and to expand their educational horizons. These differences closely mirrored the differences seen cross-nationally and therefore appear to represent differences about teacher self-efficacy across the US and Canada more than they do district-specific differences. Given that the Canadian teachers did not demonstrate higher LPI scores, it appears that teacher leadership as

measured by LPI is not indicative of self-efficacy measured by SPES and that the two have different causes. While leadership did not show association with self-efficacy, in line with work by Nir and Kranot (2006), teacher self-efficacy did correlate significantly with teacher satisfaction.

Overall, the study findings suggest that there are larger differences between *teacher empowerment* cross-nationally than there are leadership behaviors at either the administrative or the teacher level. This is intriguing because it suggests that while different school, district, state, and national education reform movements may be differentially impacting school leaders at a political, attitudinal, or structural level, the impact of these different reforms between urban schools in large districts in two Western industrialized nations did not substantially impact leadership behaviors. Self-reported leadership behaviors amongst both teachers and administrators from all of the districts were higher than previously reported LPI averages (Kouzes & Posner, 1995) and suggest that U.S. and Canadian school personnel exhibit an above average tendency toward "transformational leadership" behaviors as described by Kouzes & Posner (1995). Given the cultural similarities between the U.S. and Canada, it would be interesting to determine if this effect would be similar in high performing non Western nations. Moreover, the policies (and probably culture) at each of the levels (local, state, national) likely have an impact on how schools' teachers are positioned and their feelings of their ability to function effectively as empowered independent educators. Given that Canadian teachers overall had higher levels of empowerment, it suggests that educational policies either in Canada or at the provincial level help lay the groundwork for a school culture which is more conducive to teacher empowerment than that seen in the U.S. districts. Whether these effects on teachers are so

strongly ingrained that they negate differences in leadership or whether the two are unrelated is difficult to ascertain and would be a good topic to address in future studies.

Question Three and Four: Relationships between administrative leadership with teachers' leadership and empowerment

Although significant differences in administrator leadership were not seen in question 1, in the correlative questions 3 and 4, I looked not to compare differences cross-nationally but rather at a school level to determine whether administrator's leadership behavior as measured by their LPI subscale scores impacted either (question 3) teacher leadership behaviors as measured by LPI subscale scores or (4) teacher empowerment as measured by SPES. Spearman correlations were used to assess degree of association between each pair of factors and determine their relationships.

One important caveat for both questions 3 and 4 is that because of the uneven distribution of administrative and teacher responses, not all schools were able to be included in these analyses because there were either not administrator responses (particularly in schools where recruitment was done separately for teachers and administrators) or teacher responses (especially in schools where administrators were asked to send recruitment materials to teachers). There were a total of 26 individual schools that met the criteria of having at least one administrator and one teacher response and only 20/26 had at least five total responses and 12/26 had more than three administrator and three teacher responses.

Because more teacher than administrator responses were obtained, to ensure that all teacher responses in the 26 schools were taken into account, the correlations described include correlations of individual teacher LPI/SPES scores with their school's average administrator LPI

score. This method ensures a maximal number of data points for correlation, however may have the unintended consequence of placing too much importance on the LPI scores from a single administrator survey response in schools with only one administrator and many teachers. Alternative correlation methods (such as correlating the average of the teachers LPI/SPES versus the average LPI of the administrators across schools or across districts) would still have similar difficulties but would also have many fewer data points to use to assess correlation and were therefore not assessed. The correlations were performed both using all 26 eligible schools (in both countries) as well as in the subsets of US and Canadian schools.

For question 3, significant negative (inverse) correlations were seen between teacher's "challenge the process" with administrator's "challenge the process" and "inspire a shared vision" sub-scores. This is a very interesting finding because it suggests that when school administrators challenge their teachers or attempt to inspire them, the teachers' response is to become less likely to challenge their students to change and grow. For instance, this implies that if administrators attempt to inspire teacher to challenge current practices or share a school-wide vision/mission, then teachers are less likely to inspire their students to think creatively and "outside the box". Whether this is because teachers become conformists when administrators exhibit stronger leadership behaviors is an interesting hypothesis but not something that my study can directly address. Maybe more expectedly (and in line with research by Leech & Fulton, 2008), teacher's "encourage the heart" sub-score positively and significantly correlated with their school administrator's "model the way" sub-score indicating that administrators who provide a good example for their teachers are correlated with the teachers from their school being more likely to encourage students. Specifically, the "model the way" LPI construct

measures leaders' ability to set an example for others in the organization by aligning actions to their shared vision and doing what they say they will do (Kouzes & Posner, 1995). While significant associations, the effect size of each of these correlations was small and therefore the administrator LPI scores are not the only factors leading to the teachers' scores.

When question 3 was broken down by country, the U.S. teachers' "encourage the heart" scores were significantly positively correlated with 4 of the 5 administrator LPI sub-scores including "model the way", "challenge the process", and "enable others to act" with small to medium effect sizes. When broken down by country, the Canadian administrative LPI sub-scores did not correlate with any of the teacher LPI sub-scores. *This cross-national difference is rather striking and shows that administrative leadership in the US is much more likely to correlate with teachers' classroom leadership than Canadian administrators' leadership behaviors are to influence their teachers*. Larger sample sizes are likely needed to fully understand this difference, but it suggests that US administrators have a larger impact on their teachers' classroom leadership while some other factor (possibly structural, cultural, economic, training, or other) impacts Canadian teachers' leadership and that their school administrators' leadership does not impact what they do in the classroom in a statistically significant way. One explanation of this finding might be that Canadians have a high overall regard for public education and general support of teachers (Hargreaves & Shirley, 2012) and therefore administrative support may not be as important as in the U.S. where teachers feel less satisfied and valued (Sparks 2014).

For question 4, because there were significant differences between SPES scores both across districts and cross-nationally (as seen in question 2) and literature suggesting that

principal leadership significantly impacts teacher empowerment (Blase & Kirby,2009; Hipp, 1996), it would be attractive to suggest that administrator leadership is a leading cause for differences in teacher SPES scores. Indeed, among all eligible (>1 teacher and >1 administrator) schools, significant positive correlations were seen between teacher's "autonomy" and administrator's "challenge the process", "encourage the heart" and "inspire a shared vision" sub-scores. This indicates that while administrator leadership scores were not different cross-nationally, *their leadership does have a significant impact on the degree to which teachers feel empowered and specifically autonomous.* The negative correlation between teacher's "status" SPES sub-score and their school administrator's "model the way" sub-score is more confusing and may indicate that teachers feel less empowered and important if their school leaders exhibit strong leadership behaviors.

Cross-nationally, the U.S. teachers' "self-efficacy" scores were significantly positively correlated with administrator "inspire a shared vision" LPI sub-score with a medium effect size. Like the overall analysis, significant positive correlations were seen between U.S. teacher's "autonomy" and administrator's "challenge the process", "encourage the heart" and "inspire a shared vision" sub-scores, these showed small to medium effect sizes. Because these were not seen as significant in the Canadian analysis, this indicates *that most of the association that was seen in the overall correlation came from the US schools*. This difference—where US administrator LPI impacted US teachers' SPES and empowerment more than Canadians—is similar to the contrast seen with teacher LPI in question 3.

Canadian teacher's "status" SPES sub-score negatively correlated with "model the way", "enable others to act" and "inspire a shared vision" administrative LPI sub-scores means that the

Canadian response drove the overall negative association between teacher's "status" SPES sub-score and their school administrator's "model the way" sub-score. It is also somewhat perplexing that teacher SPES "decision-making" negatively correlated with the administrator's "inspire a shared vision" sub-score meaning that when Canadian administrators tried to inspire their teachers, the teachers from their schools felt like they had less ability to make and impact decisions in their school.

The relationships between teachers' leadership and empowerment with administrator leadership are not surprising based on literature (Nir & Kranot, 2006; Bogler, 2001) suggesting the importance of school leadership on promoting teachers' self-efficacy, effectiveness, and students' academic achievement. However, the individual types of administrator leadership which impact teachers were telling. While administrators' "model the way" score impacted teachers' leadership behaviors, and particularly "encourage the heart", their "challenge the process", "encourage the heart" and "inspire a shared vision" scores correlated with teacher autonomy. This suggests that different exemplary leadership behaviors in administrators have different effects on teachers and change the school culture differently: principals who set a good example are more likely to create a culture of teacher leadership while those who are motivational and encouraging are more likely to inspire teachers to feel independent and effective. This has important implications for the types of leadership that schools may wish to employ. If schools wish to have a culture of leadership including a more distributed leadership with teacher leaders, then working on leading by example would be a key aspect for administrators whereas other actions such as giving teachers praise, developing a common set of values, and innovating would have more impact on creating teacher satisfaction, empowerment,

and autonomy but would not necessarily impact teacher leadership. The concepts of both transformational and distributed leadership have become increasingly important in the global education reform context, but "the intensifying pressures of globalization will undoubtedly necessitate the evolution of existing theories along with the gradual emergence of a wide array of new and relevant leadership paradigms" (Litz, 2011, p. 58).

This study builds directly on previous studies (Richardson, Flanigan, Lane, & Keaster, 1992; Leech & Fulton, 2008; Leech, Smith, Green, & Fulton, 2003) of leadership which address the effects of administrators' leadership on schools and on teachers. However, in addition to exploring whether administrators demonstrated leadership and how that demonstrated leadership impacted teachers, this study addresses multiple gaps in the literature. First, this represents the first study to directly compare administrator and/or teacher leadership across U.S. and Canadian urban secondary schools. While case studies (Lin & Shi, 2014; Skerrett, 2010) have compared U.S. and Canadian districts in educational outcomes, this study specifically addresses leadership both across two large districts and also cross-nationally. Furthermore, by doing correlational analysis in questions three and four, my study also extends the findings of multiple scholars who have previously demonstrated the importance of promoting teacher leadership (Moller et al., 2001; York-Barr & Duke, 2004), self-efficacy (Wall & Rinehart, 1998), satisfaction (Bogler, 2001; Perrachione, Rosser, & Peterson, 2008), and empowerment (Sweetland & Hoy, 2000). While these studies have previously shown correlations between leadership and other outcomes, my present study clarifies these differences in the context of urban U.S. and Canadian secondary schools. While the differences observed are no necessarily causal of the differences in educational outcomes seen between these countries on international tests where Canadian

students outperform their US counterparts (OECD, 2011), however the correlations seen do suggest that differences in teacher leadership and empowerment exist and may play a role in such outcomes.

Furthermore, this study utilizes the framework of leadership set out by Kouzes and Posner (1995, 2002) which defines 5 key leadership behaviors as well as the concept and conceptual structures of teacher empowerment defined by Short and Rinehart (1992) to assess teacher and administrator leadership. While the way in which these individual tools are implemented is not novel, their being assessed together and the comparison of the LPI sub-scales with the SPES sub-scales has not previously been reported. The fact that this study shows that multiple administrator LPI scores correlate with teacher SPES scores expands the notion that administrator leadership impacts teachers and provides new insights into some of the nuanced ways in which these constructs may interact at the school level. Given the multiple lines of evidence (Leithwood & Jantzi, 2005; O'Donnell & White, 2008) suggesting the importance of leadership in defining school culture, scholars such as Kouzes and Posner (2002), Leithwood and Jantzi, (2000) York-Barr and Duke (2004) would likely not be surprised that administrator leadership behaviors impact teachers' satisfaction and empowerment. Given previous literature (Chin, 2007; Robinson, Lloyd, & Rowe, 2008) suggesting administrative leadership's fundamental role in predicting educational performance, the lack of differences in administrative leadership seen between US and Canadian districts might confuse these authors given the knowledge that Canadian students outperform US students on international assessments. However, given that there were substantial differences in teacher leadership cross-nationally, I

contend that authors like Kouzes and Posner would likely interpret the study as showing the importance of teacher leadership in this cross-national context.

With regard to district administrative leadership, based on historical data, the average LPI scores for both U.S. and Canadian respondents suggested transformational leadership. When normalized to compare to historic LPI responses drawn from thousands of respondents as described by Kouzes and Posner (1995), administrators demonstrated scores of 53.3, 51.8, 53.4, 51.4, and 49.6 for "model the way", "challenge the process", "enable others to act", "encourage the heart", and "inspire a shared vision" compared to historic leader averages of 47.0, 43.9, 48.7, 43.8, 40.6. Some researchers have used a composite of these scores as a measure of transformational leadership; with a mean score of 51.9 for samples of sampled administrators vs 44.8 historically, U.S. and Canadian administrators self reported high levels of transformational leadership. While this study did not assess measures of other types of leadership (i.e. transactional leadership), given previous literature (Bogler, 2001) suggesting that leaders primarily utilize one of these two styles, it is apparent that the administrators in this study were more likely to use transformational rather than transactional leadership.

This study also utilized the LPI instrument to assess teacher leadership, and again, teachers demonstrated higher average scores than historically reported levels of transformational leadership with scores of 52.2, 50.4, 55.2, and 53.4 for "model the way", "challenge the process", "enable others to act", and "encourage the heart" compared to historic leader averages of 47.0, 43.9, 48.7, and 43.8. Using the composite average, teachers from this study averaged 52.8 compared to a historical average of 45.9. While this would certainly characterize these teachers more as transformational rather than transactional leaders in their classrooms, it is not

fully able to address the question of whether the teachers are what York Barr and Duke (2004) term 'instructional' or 'organizational' leaders. While not a full assessment of these concepts, one can look to the SPES areas of Decision-making, Status, and Autonomy as surrogates for how much teachers view themselves as organizational leaders. Given that scores were high (average >8/10) in each of these categories, one can conclude that a majority of teachers viewed themselves as organizational leaders in addition to their roles as instructional leaders. Of note, Canadian teachers had significantly higher scores for autonomy and decision-making compared to U.S. teachers and may therefore have more of a role in their schools' organizational leadership. In the sense that teachers were given more responsibility and felt more autonomous, one could speculate that the Canadian schools' administrators may have demonstrated higher levels of distributed leadership, but triangulating this from teachers' empowerment data is inductive and a different leadership survey (such as the distributed leadership inventory, Hulpia, Devos & Rosseel, 2009) would need to be administered to address this question directly.

Limitations

Limitations Inherent In Survey Research

Survey research is inherently limited in a number of ways. First, survey research is susceptible bias such as volunteer/respondent bias, to variable response rates, and to sampling error (the surveys are not sent to a population representative of the populations as a whole) (Visser, Krosnick, & Lavrakas, 2000). Survey respondent honesty is another issue. While most respondents will answer most kinds of surveys honestly, the anonymous nature of surveys enables respondents to provide false responses (Rogers & Richarme, 2009). Some measures can

be taken to reduce dishonest surveys, but survey researchers do depend on the general honesty of their participants. Furthermore, Rogers and Richarme (2009) note that in general, the issue of honest responses is considered less substantial for focused or solicited surveys such as those in this study than for those available to the general public.

Survey research in general also limits the range of responses that can be obtained. Because respondents are not able to explain their answers, there may exist hidden confounding variables that explain study results but which were not asked and therefore cannot be assessed. Quantitative results based on survey responses can be used to clearly explain that there are differences, correlations, and associations, but the causality of these statistical findings generally cannot be ascertained. The use of validated quantitative instruments assures that the questions asked deliver reliable responses, but without field observation or qualitative interviews, there still remains a possibility that important factors are missed.

Furthermore, survey self-reporting can only provide information about past or likely actions. Responses may not describe how people will actually act in a given situation. It is important to note that self-reported leadership behaviors as ascertained by surveys may not correlate entirely with actual leadership behaviors. According to Azjen's (1985) theory of planned behavior, attitudes toward a behavior, subjective norms, and perceived behavioral control, work together to shape an individual's behavioral intentions and behaviors. While these latter 2 factors (subjective norms and perceived behavioral control) are potentially relevant and are not assessed with the survey instruments in this study, it should be noted that attitudes/intentions are often good predictors of specific behaviors and that this correlation is the basis of many contemporary theories of human behavior (Ajzen & Fishbein, 2005).

Limitations Due To Survey Instrument

The LPI and SPES instruments have been shown to be valid and reproducible (Kouzes & Posner,2005; Short & Rinehart, 2002) and multiple previous studies have used the LPI instrument to assess principal/administrative leadership (Leech & Fulton 2002; Taylor, Martin, Hutchinson, & Jinks,2007). The teacher survey utilized in this study is a unique combination of previously validated questions from the LPI and the SPES instruments but to ensure the survey's brevity, it does not comprise the entirety of either survey in full. The validity of the individual questions used, therefore should not be problematic, but the subscale scores / factors identified in the teacher survey are drawn from a smaller number of items than from the original survey. Therefore, some caution must be taken in interpreting results from the teacher survey. Despite this potential concern, the high correlations across questions within the same subscale argue that subscale internal validity remains and that interpretations are not likely to be substantially effected.

Limitations Based On Sampling And District-Specific Methodologic Variability

In the present study, one of the major limitations was the variability across US and Canadian districts, both in terms of the types of schools and districts as well as in the way that districts would allow for research to occur. After careful evaluation of multiple districts across factors such as size, diversity, school types, and history, two districts were initially selected as comparators, but these districts did not approve the research study.

In order to ensure this study was able to occur, I applied to fifteen US and nine Canadian school districts/boards (chosen on the basis of size, demographics, and research review process)

over the course of six months. Each district application required different information and modification of various aspects of the research. While some of this variability was small (i.e. changing the language in the principal recruitment letter), other changes made the way that recruitment occurred or the study took place fundamentally different. For example, some districts (U1, C2) allowed me to contact teachers directly via email while most districts required that the school's principal contact the teachers to ask them to participate. This seemingly minor difference was important because while some principals (i.e. those in district C1) solicited their teachers' responses, there were other districts (U3, U4, C4) where poor teacher response rates may have resulted from lack of principals' sending the recruitment materials to teachers. Another major difference was one district (C4) did not allow me to collect the name of the school in my survey and therefore participants from this district were included in the cross-national analysis of questions 1,2, and 5 but had to be excluded from the question 3 and 4 analysis because school level matching could not be done.

Limitations From Low Response Rates And Limited District Sample Size

Another important limitation of this study was sample size and response rate. While response rates remained lower than expected (1-10%) and lower than reported in similar studies (Starcher 2006; Garity 2011), eventual sample sizes were sufficiently large to allow correlational analyses (including analysis of administrator LPI with teacher LPI and with teacher SPES) with results having confidence intervals sufficient to make statistical conclusions. Furthermore, the low response rates are not inconsistent with previous studies using online surveys which frequently report rates substantially lower than paper surveys (Nulty, 2008).

To increase response rates, multiple strategies were implemented. In addition to initial recruitment emails which highlighted the importance, benefit to schools/districts, and brevity of the study, all approved potential participants received a two-week follow-up and a 'last-call' email reminding them to participate in the survey. In addition, I conducted several meetings with district research personnel and eventually gained additional approval to have the district include recruitment flyers in district newsletters. In one district, I contacted the teacher's union and had them send an email advertising the study. Some districts asked for examples of "the direct benefit to district" and in addition to providing detailed study results, I offered to provide free consultations to district personnel/administrators to go over their particular schools' results and discuss potential implications and actionable items to help improve school improve on any areas seen as deficient.

Beyond these types of measures, another type of participant recruitment which is frequently used to improve response rate is participant compensation (Visser, Krosnick, & Lavrakas, 2000). Without study funding, compensation was not available but having participants entered into a drawing (for a gift card, iPod, etc.) may have been possible. However, this was deemed coercive and not allowed by multiple districts and school boards and therefore was not utilized.

Some issues related to response rate are structural and could only be addressed by performing the study in another way (i.e. on-site distributed paper surveys, etc.). For example, one district's response rate was initially zero because recruitment emails were not received by staff after they were blocked by a district-wide spam filter. After this was addressed, school personnel had to send all recruitment materials. While this led to some improvement in response rate, the lack of

personal follow-up which I did with other districts was lacking and overall response rates remained lower than other districts.

Another major issue was that districts did not provide a list of teachers or administrators or their email addresses and information frequently had to be obtained via websites and other publically available data. In addition to being tedious and inefficient, this method was limited by the fact that many district websites were out of date with incorrect staff information did not have email addresses or only had limited email contact information. While I was able to use programming procedures to extract relevant emails, many emails (upwards of 5-15% per district) still "bounced", and many others may have been incorrect leading to lower response rates than would have been seen had the study been sent to all appropriate/eligible staff. I contacted district personnel from all eight participating districts asking for relevant contact information, but none provided any more than names of principals from eligible/approved schools.

Implications

Implications for Education

Based on the results of this study, administrator leadership qualities may not vary cross-nationally, but they have important impacts on teachers and both their classroom leadership and their empowerment. Administrators' "model the way" score correlated positively with both teachers' "encourage the heart" sub-score as well as their "status" score. This indicates that administrators who provide a good example for their teachers led the teachers in their school both to be more likely to encourage students as well as to feel respected and important in the school. Similarly, administrator's scores on "challenge the process", "encourage the heart" and "inspire a shared vision" sub-scores all correlated with teacher autonomy. These results indicate

that school leaders who provide high levels of these leadership qualities not only inspire their teachers to lead, but also to feel more independent and important. These aspects are crucial in developing a positive school culture and therefore administrators looking to improve staff morale, cohesion, and leadership would benefit from focusing on some of the specific behaviors which the LPI survey focuses on such as praising teachers for a job well done, following through on promises, and setting a personal example.

Intriguingly, there were cross-national differences in not only in teacher leadership and empowerment, but also in the degree to which administrator leadership impacted teachers. In the US, there were significant differences. For example, U.S. teachers' "encourage the heart" scores were significantly positively correlated with 4 out of 5 administrator LPI sub-scales while the Canadian administrative LPI sub-scores did not correlate with any of the teacher LPI sub-scores. This reveals that administrative leadership in the US is much more likely to lead to changes in teachers' classroom leadership. Therefore, while all administrators would likely benefit from improving their leadership, the impact that improved leadership had was more evident in US schools and therefore US administrators had more ability to impact their teachers' classrooms by improving their leadership behaviors than did Canadian administrators.

Another important implication is that US participants were more likely to attain a higher degree of formal education compared to Canadians while Canadian participants were more likely to have more years of professional experience. These factors were shown to be important contributors to some of the LPI and SPES outcomes and therefore are of interest. Given the lower satisfaction scores in the US compared to Canada, it would be interesting to see what

extent post-baccalaureate education impacts teachers and whether leadership or other factors are responsible for Canadians' tendency to stay in high school jobs longer than their US colleagues.

Implications For Participant Districts And Schools

For the individual participant schools and districts in the study, the implications lie in the notion of learning from others as an imperative for schools or districts which underperform in any of the leadership or empowerment sub-scales. As part of my district reporting process, I will provide district (and in the case of districts U1 and C1, school) level reports which will detail not only the overall study results but also district/school responses and how they compare to other districts. These data will provide both raw data (tables) as well as with interpretation of districts' relative areas of strength / areas which need improvement with respect to the study's outcome measures (LPI, SPES). I hope that district personnel will be able to use these reports to assess how administrator leadership can be improved in the schools where the study was done so that this study can foster better working environments and incrementally move toward improving conditions supportive of teacher leadership, satisfaction, and empowerment. Some districts will require me to present and discuss study results with district/school personnel and provide further guidance into how these data can be transformed into actionable items at the district and the school levels.

Implications For School Administrators

The results of this study indicate that the behaviors of school administrators and the leadership which they demonstrate have important implications not only for teachers in their

school but for school culture and educational outcomes. This implies that the "transformational" leadership behaviors indicated by school leaders' responses have a trickle-down effect and can improve schools at multiple different levels. Therefore, while principals' and administrators' formal training is often focused on improving student scores on standardized tests, administrators are not formally trained in leadership practices or effectively evaluated on the basis of how they impact the school. Indeed, "state and district evaluations do not reflect existing principal standards or proven practices, and many principal evaluation instruments are neither technically sound nor useful for improving principal performance" (NAESP, 2016) meaning that school leaders aren't being judged on the behaviors that have the most potential to impact the school. Instead, it would be advantageous if administrators were judged on their leadership behaviors and their staff's responsiveness to their leadership. If a principal knew they were deficient in encouraging their staff, for example, this is a far more remediable problem - because this is more within an administrators' direct control than low test scores. In other words, this research implies that if changes are made in a specific administrative leadership domain, there are likely to be tangible educational benefits to the school in the form of more effective teacher leadership.

Implications For Teachers

This work speaks to the importance not only of teacher leadership, but also to the primacy of teacher empowerment and teacher satisfaction. Teacher attrition remains an issue in the U.S. (as evidenced by the significantly lower years of experience in this study) and is one area which needs to be addressed (Clandinin et al, 2013). In the global educational sphere, teachers are frequently impacted by the school leaders that they encounter. In line with literature

from the International Successful School Principals Project (ISSPP), a 20 country network of researchers with over 100 case studies of principals who have built and sustained success in different international contexts and sectors (Day & Leithwood, 2007; Moos et al. 2011), this study suggests that school administrators play a vital role in establishing the "conditions, structures, cultures and climate for professional learning and development in their schools" (Day, 2013, p. 32). By better understanding both the local and national contexts which impact teachers and then helping administrators create a school culture to best address their issues, this study suggests that leadership can have a major impact on teacher satisfaction and empowerment, and that can lead to improvements in educational outcomes.

Implications For National And Global Educational Policymakers

Sahlberg (2011) speaks of global education reform as stuck in a mire of so-called "GERM" education which relies on the basic tenets of standardization, focus on core subjects, attempts to find low-risk ways to achieve learning goals, corporate management models, and test-based accountability. Sahlberg is correct that test-based models are insufficient to impact learning in the most important ways. Instead, he suggests that teacher and principal education/training and satisfaction are key components to a successful educational reform program. The lesson from this study with regards to level of education and experience is somewhat interesting because the more highly educated U.S. teachers tended to be less satisfied and feel less empowered than their Canadian counterparts. While this does not directly contradict what Sahlberg discusses with respect to the importance of teachers in Finland, it suggests that other factors must be taken into account when policymakers are trying to improve

education at the state or national level. Because U.S. teachers are paid less and respected less than teachers in other nations (Dolton & Marcenaro-Gutierrez, 2013), and because they tend to have lower professional autonomy and less time to collaborate as professionals (Darling-Hammond, 2011), they also have a higher tendency to burnout. Therefore, while policies which encourage additional teacher training may be beneficial, they should only be implemented after policies which ensure teacher self-efficacy and satisfaction. This may mean that national reform agendas should focus more on programs to better compensate, support, and motivate teachers (including promoting principal transformational leadership and distributed leadership within schools) than on level of education and certainly more so than test-based accountability. In line with the Harvard University "Education for the 21st Century" report (Reimers & Kanter, 2014), I would advocate a policy agenda which prioritizes bold educational reforms including (1) creating an inspiring sense of purpose, (2) preparing educational leaders to drive needed change using leadership skills, (3) recruiting and retaining better teachers, and (4) involving local communities in educational reform efforts. Such an agenda would not only move reform away from GERM tactics, but would also integrate administrators and teachers as vital and invested leaders in the move toward educational improvements.

Future Directions

Expanding The Study Population

The results of this study indicate differences between two individual districts (U1 and C1) and show differences across four districts. However, the generalizability of the study to the US and Canada as a whole would require a larger number of districts, schools, and participants. Therefore, the extension of these results performed in additional districts would ensure that the

current results are valid as a cross-national study and determine if the differences observed in the present study are consistent across multiple districts in each country. This could be also be done to ensure an increase in statistical power to >80%.

Performing The Research Closer To The Participants

Surveys from external researchers whom study participants do not know and will not interact with significantly are frequently ignored or are not taken seriously (Lorenc et al, 2013). One way to avoid this situation and to improve the relevancy of research to individual respondents is to make the study more locally relevant and contextualized (Kinder Institute for Urban Resarch, 2012). Whereas the current research focused on cross-district and cross-national implications, individual districts and participants are more likely to be concerned about the local implications of the research. For that reason, some scholars (e.g., Cochran-Smith & Lytle, 2009) advocate performing the study in conjunction/collaboration with local school board researchers, teachers unions, and other local stakeholders so that the study can benefit from increased buy-in from principals and other district personnel. By performing the work as an external researcher and administering surveys electronically, I faced barriers to achieving optimal response rates. By working more closely with local districts, future studies could not only add additional local contexts to the study, but also address the low response rate.

Mixed Methods

One way to improve our understanding of the current research would be to perform additional qualitative analyses or a mixed methods study in order to better assess some of the reasons why the differences observed are present. Because quantitative studies are limited in

their ability to assess causality and can only report answers that have already been pre-defined, asking more open-ended questions and hearing broader perspectives from teachers and administrators would likely shed additional insights on what differences occur both structurally and behaviorally in schools across the US and Canada. While such a study would have to be smaller in scope than the present study, interviews and observations of practitioners might make clear some of the reasons that teachers in Canada feel more empowered or why US administrators' leadership is more likely than Canadian administrators' leadership to influence teacher leadership. While the differences are fascinating in this study, hearing the reasons that the differences exist would make it easier to understand how to modify schools and classrooms to maximize teacher leadership and empowerment.

The importance of state/provincial and local contexts must also be considered in the interpretation of my results and would be something that could be inherently better addressed using qualitative methods. Only by being able to explore the state/local political and educational policies in schools being studied and their implementation at the school level would it become fully evident how these potential differences would impact study findings. Given a substantial literature suggesting that local contexts are important for a full understanding of education across schools, the ability to fully characterize schools' local context on participants' responses would yield important insights into how the results may be a reflection not only of national differences, but also of state and local contexts.

My original dissertation proposal consisted of a mixed methodology in which I planned to follow up the results of a quantitative study with a more in-depth qualitative study (based on observations and interviews of participants during travel to two districts) of volunteer

participants from the quantitative portion of the study. Such a design would have allowed me to dig deeper into the reasons behind the differences observed in a quantitative part (like the cross-district study within this work), collect on-site documents that guide school administrators and teachers, and use the open-ended responses and site observations (of the teaching and learning environment, and leadership practices) as a way to triangulate why some of the observed differences may exist. Unfortunately, for logistical reasons (time, money, departmental support), a study of such scope was not feasible in the context of my dissertation. In line with these original intentions, such a future mixed-methods study would yield significant insights into why quantitative results occur, beyond the scope of the current work.

There were some cross-national differences observed that were strikingly statistically significant but which did not correlate with administrative leadership. Teacher satisfaction was the best example of this and represented the single largest difference between teachers either between districts U1 and C1 or cross-nationally. Given the lack of association of satisfaction with any administrative LPI category, it is likely that differences in the educational systems in Canada lead to higher-level job satisfaction as opposed to things that the administrators are doing. Future research will have to specifically address causes of teacher satisfaction and determine what things are happening in Canadian schools that lead to higher satisfaction and if any of these things might be applicable to US schools to improve overall teacher happiness. Given the high burnout rate in US education (Bousquet, 2012), this could be beneficial not only to teachers but also to US policymakers and education in general.

Another aspect that future qualitative/mixed methodology studies would need to address is the role of ethnic and economic diversity in the study results. While the demographic

assessments and corrections done in question 5 give important insights, given the important differences in disadvantaged teachers' and students' classroom experience (Payne 2008), hearing stories of how administrative and teacher leadership impact school culture differ between schools of different socioeconomic status and ethnic makeups would also add important additional perspectives. While there is increasing scholarship about the different types of educational disparities that occur in poor urban schools, there is a lack of literature suggesting how school leadership in these types of settings can overcome their unique goals, and the voices of successful school leaders and their utilization of leadership behaviors to inspire teachers is sorely missing. Triangulating the stories of leaders at high performing urban schools with the empirical research done in this study has the ability to yield important insights into how leadership behaviors can transform schools, improve culture, morale, and teacher empowerment—particularly in schools, like those in economically depressed areas, where more effective leadership may be most impactful and most needed.

Use More Extensive / Longer Surveys

Another way to extend the results of this study would be to perform more in-depth surveys to more fully understand principal leadership than the LPI which is a more generalized leadership inventory. Principal leadership has traditionally been measured using the Principal Instructional Management Rating Scale (Hollinger, 1983), but recent research has shown that more focus is needed on standards that are linked to behaviors which improve student achievement. The ISLLC (Interstate School Leaders Licensure Consortium) standards, which were adopted in 2008, define such standards and a recently designed tool, the Vanderbilt

Assessment of Leadership in Education (VAL-ED) has been shown to meet ISLLC standards (Porter et al., 2010). While VAL-ED has numerous advantages (Wallace Foundation, 2009) including its being tied to leadership actions that promote improved achievement, it is a 25 minutes, 72 questions format which requires "360 degree" leadership evaluations by all school teachers and staff. A similar assessment, CALL (Comprehensive Assessment of Leadership for Learning) has been under study since 2011 (Halverson, Kelley, & Shaw, 2014), is 100 questions long and takes 40 minutes to complete and assesses principals' distributed leadership qualities as assessed by their staff. While these surveys were too time-intensive to perform in a large number of schools cross-nationally, their use in a small subset of schools from the present study could yield additional insights into school culture and how administrative leadership impacts teachers and staff.

Urban vs Suburban vs Rural

The current study focuses only on urban schools that have a very different set of characteristics than national school systems as a whole. To better address whether there are "cross national" differences, this study would need to be performed in suburban, and rural districts as well to determine if the results are specific to urban districts or whether differences observed are truly different at the national level. This type of study would also have the benefit of assessing differences in leadership across these different types of schools within each country, and whether administrative leadership in different school settings leads to similar effects on staff.

Performing a Multi-National Study

As indicated in the review of literature, the US lags a number of other nations on objective cross-national assessments such as PISA and TIMSS. While Canada represents the strongest, most similar and a logical comparison, a larger multi-national study looking at leadership using LPI and SPES surveys in top performing nations such as Finland, Singapore, and Japan would be able to better determine the degree to which administrative leadership impacts school achievement. The current study also only gives voice to Western schools and the leadership behaviors and their impact may be significantly different in, for example, Eastern or developing countries; only a comprehensive multi-country study could address this gap and give voice to multiple different educational perspectives across the globe. The surveys would need to be translated into a number of different languages in order to do this, which can sometimes cause difficulty in interpretation of the results. If language posed too substantial a barrier, at least extending the study to other English speaking nations such as Great Britain and Australia would still yield insights into how leadership varies across the globe.

F. CONCLUSION

Given the increasing evidence that U.S. schools lag behind many of their international counterparts and growing literature suggesting that leadership is essential to ensure that schools succeed, this work assessed differences in school leadership behaviors in the U.S. and Canada and correlated these behaviors with other important educational outcomes.

This study demonstrates that there are significant differences between teacher leadership and empowerment both across districts and in a cross-national study of eight districts but that similar differences do not exist across administrative leadership. Indeed, the data show that Canadian teachers had significantly higher 'job satisfaction' compared to U.S. teachers ($p<0.001$), more overall sense of empowerment as measured by the composite SPES ($p<0.001$), and significantly higher scores in the SPES sub-scale categories of "professional growth" ($p=0.001$), "autonomy" ($p<0.001$) and "decision-making" ($p=0.028$) where by comparison there were no statistically significant differences among any of the administrator LPI sub-categories.

Administrator leadership, however, was significantly correlated with teacher leadership and empowerment at the school level and therefore is a vital factor in relevant teacher outcomes. Specifically, I found significant positive correlations between teacher's "autonomy" and administrator's "challenge the process", "encourage the heart" and "inspire a shared vision" LPI sub-scores. Conversely, teacher's "status" SPES sub-score was significantly negatively correlated with their school administrator's "model the way" sub-score. This suggests that when school leaders wish to improve teachers' job satisfaction and performance, it is important for them to consider how they engage their staff through words of encouragement, working toward a

school mission, and challenging the status quo in order to impact the autonomy of the teachers in their school.

This study suggests that there are significant differences in teacher leadership and teacher empowerment between two large districts (one in the US and one in Canada) and cross-nationally across the eight participant districts (four in the US and four in Canada). Through careful study design and application of inferential statistics, these differences suggest that there are cross-national differences in similar urban districts. While this study indicated that some of the differences in leadership behaviors are associated with administrator leadership, such associations did not explain all of the variance in the teachers' survey. Demographics, particularly years of experience, level of education, and ethnicity were different across districts and cross-nationally, but analysis showed that these differences were not the main reason for the differences in leadership and empowerment. Teachers in the four US urban school districts were less experienced, more educated, and more diverse than their Canadian counterparts. They displayed similar leadership behaviors and self-efficacy, but had lower levels of satisfaction, autonomy, and empowerment. An executive summary describing all study results is outlined in Table 32.

This work builds critical new knowledge in the areas of educational leadership and international education by extending scholarship about school leadership to assess impact on teacher empowerment in a cross-district and cross-national context. It also lays the foundation for ongoing scholarship that can further elucidate some of the causes for the differences and correlations observed and help practitioners turn this knowledge into actionable items that can be used to improve leadership and educational outcomes in schools. In particular, it indicates the

utility of performing additional research, particularly qualitative/mixed methodologies research, which can further pinpoint the root causes of the differences and correlations illuminated by the present study. Such work will give impetus for how to address the differences uncovered in this study with actionable items for the schools and districts studied and give insights into how such strategies might have cross-national applicability.

Table 23. Executive Summary of Study Results

Research focus	Cross-district (U1/C1) results	Cross-national (8 district) results
1. Administrator leadership	❖ No cross-district differences	❖ No cross-national differences
2. Teacher leadership, satisfaction, and empowerment	❖ U1 teachers had higher "challenge the process" scores ❖ C1 teachers had higher levels of "satisfaction", "empowerment", "professional growth", "autonomy", and "decision-making"	❖ U1 teachers had higher "challenge the process" scores ❖ C1 teachers had higher levels of "satisfaction", "empowerment", "professional growth", "autonomy", and "decision-making"
3. Relationships exist between administrative and teacher leadership practices		❖ Negative correlations were seen between teacher's "challenge the process" with administrator's "challenge the process" and "inspire a shared vision" sub-scores. ❖ Teacher's "encourage the heart" sub-score positively correlated with their school administrator's "model the way" sub-score. ❖ U.S. teachers' "encourage the heart" scores were significantly positively correlated with 4/5 administrator LPI sub-scores including "model the way", "challenge the process", and "enable others to act" scores ❖ Canadian administrator scores did not impact teacher scores
4. Relationships exist between administrative leadership practices and teacher empowerment		❖ Positive correlations were seen between teacher's "autonomy" and administrator's "challenge the process", "encourage the heart" and "inspire a shared vision" sub-scores. ❖ Teacher's "status" SPES sub-score negatively correlated with their school administrator's "model the way" sub-score. ❖ U.S. teachers' "self-efficacy" scores were positively correlated with administrator "inspire a shared vision" sub-score

G. SUPPLEMENTAL TABLES

Supplemental Table 1. Cross-District Analysis of Administrator LPI Items. Mann-Whitney tests were performed for all 30 individual items from the administrator LPI across district U1 and district C1. Statistically significant differences were shown for items 12 and 19 ($p<0.05$).

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-1. I set a personal example of what I expect of teachers. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	1.000[1]	Retain the null hypothesis.
2	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-2. I follow through on the promises and commitments that I make. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.060[1]	Retain the null hypothesis.
3	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-3. I ask for feedback on how my actions affect teachers' performance. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.758[1]	Retain the null hypothesis.
4	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-4. I spend time and energy making certain that the teachers I work with adhere to the principles and standards we have agreed on. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.718[1]	Retain the null hypothesis.
5	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-5. I build consensus around a common set of values for running our organization. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.904[1]	Retain the null hypothesis.
6	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-6. I am clear about my philosophy of leadership. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.289[1]	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[1] Exact significance is displayed for this test.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
7	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-7. I challenge teachers to try out new and innovative ways to do their work. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	1.000[1]	Retain the null hypothesis.
8	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-8. I search outside the formal boundaries of my organization for innovative ways to improve what we do. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.142[1]	Retain the null hypothesis.
9	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-9. I ask "what can we learn?" when things don't go as expected. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.081[1]	Retain the null hypothesis.
10	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-10. I seek out challenging opportunities that test my own skills and abilities. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.445[1]	Retain the null hypothesis.
11	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-11. I make certain that we set achievable goals, make concrete plans, and establish measurable milestones for the projects and programs that we work on. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.738[1]	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[1] Exact significance is displayed for this test.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
12	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-12. I experiment and take risks, even when there is a chance of failure. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.035[1]	Reject the null hypothesis.
13	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-13. I actively listen to diverse points of view. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.620[1]	Retain the null hypothesis.
14	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-14. I treat teachers with dignity and respect. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	1.000[1]	Retain the null hypothesis.
15	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-15. I ensure that teachers grow in their job by learning new skills and developing themselves. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.060[1]	Retain the null hypothesis.
16	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-16. I develop cooperative relationships among the teachers I work with. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.072[1]	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[1]Exact significance is displayed for this test.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
17	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-17. I support the decisions that teachers make on their own. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.968[1]	Retain the null hypothesis.
18	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-18. I give teachers a great deal of freedom and choice in deciding how to do their work. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.221[1]	Retain the null hypothesis.
19	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-19. I praise teachers for a job well done. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.030[1]	Reject the null hypothesis.
20	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-20. I make it a point to let teachers know about my confidence in their abilities. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.445[1]	Retain the null hypothesis.
21	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-21. I find ways to celebrate accomplishments. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.414[1]	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[1]Exact significance is displayed for this test.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
22	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-22. I make sure that teachers are creatively rewarded for their contributions to the success of our projects. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.277[1]	Retain the null hypothesis.
23	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-23. I publicly recognize teachers who exemplify commitment to shared value. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.904[1]	Retain the null hypothesis.
24	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-24. I give the members of the team lots of appreciation and support for their contributions. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.242[1]	Retain the null hypothesis.
25	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-25. I talk about future trends that will influence how our work gets done. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.904[1]	Retain the null hypothesis.
26	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-26. I describe a compelling image of what our future could be like. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.904[1]	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[1]Exact significance is displayed for this test.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
27	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-27. I appeal to teachers to share an exciting dream of the future. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.947[1]	Retain the null hypothesis.
28	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-28. I show teachers how their long-term interests can be realized by enlisting in a common vision. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.678[1]	Retain the null hypothesis.
29	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-29. I paint the "big picture" of what we aspire to accomplish. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.883[1]	Retain the null hypothesis.
30	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-30. I speak with genuine conviction about the higher meaning and purpose of our work is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.799[1]	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[1]Exact significance is displayed for this test.

Supplemental table 2. Cross-National Comparisons of Individual Administrator LPI Item Tests. Mann-Whitney tests were performed for all 30 individual items from the administrator LPI across all participant US and Canadian district administrators. Statistically significant differences were shown for items 2, 3, and 19 ($p<0.05$). Items 2 and 19 showed higher mean ranks from Canadian administrators while US administrators had significantly higher ranks for item 3 (full data not shown).

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-1. I set a personal example of what I expect of teachers. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.981	Retain the null hypothesis.
2	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-2. I follow through on the promises and commitments that I make. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.036	Reject the null hypothesis.
3	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-3. I ask for feedback on how my actions affect teachers' performance. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.044	Reject the null hypothesis.
4	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-4. I spend time and energy making certain that the teachers I work with adhere to the principles and standards we have agreed on. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.812	Retain the null hypothesis.
5	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-5. I build consensus around a common set of values for running our organization. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.826	Retain the null hypothesis.
6	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-6. I am clear about my philosophy of leadership. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.075	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
7	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-7. I challenge teachers to try out new and innovative ways to do their work. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.147	Retain the null hypothesis.
8	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-8. I search outside the formal boundaries of my organization for innovative ways to improve what we do. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.480	Retain the null hypothesis.
9	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-9. I ask "what can we learn?" when things don't go as expected. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.915	Retain the null hypothesis.
10	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-10. I seek out challenging opportunities that test my own skills and abilities. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.977	Retain the null hypothesis.
11	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-11. I make certain that we set achievable goals, make concrete plans, and establish measurable milestones for the projects and programs that we work on. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.467	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
12	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-12. I experiment and take risks, even when there is a chance of failure. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.151	Retain the null hypothesis.
13	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-13. I actively listen to diverse points of view. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.316	Retain the null hypothesis.
14	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-14. I treat teachers with dignity and respect. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.848	Retain the null hypothesis.
15	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-15. I ensure that teachers grow in their job by learning new skills and developing themselves. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.823	Retain the null hypothesis.
16	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-16. I develop cooperative relationships among the teachers I work with. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.372	Retain the null hypothesis.
17	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-17. I support the decisions that teachers make on their own. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.227	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
18	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-18. I give teachers a great deal of freedom and choice in deciding how to do their work. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.472	Retain the null hypothesis.
19	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-19. I praise teachers for a job well done. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.009	Reject the null hypothesis.
20	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-20. I make it a point to let teachers know about my confidence in their abilities. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.366	Retain the null hypothesis.
21	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-21. I find ways to celebrate accomplishments. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.622	Retain the null hypothesis.
22	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-22. I make sure that teachers are creatively rewarded for their contributions to the success of our projects. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.351	Retain the null hypothesis.
23	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-23. I publicly recognize teachers who exemplify commitment to shared value. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.531	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
24	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-24. I give the members of the team lots of appreciation and support for their contributions. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.724	Retain the null hypothesis.
25	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-25. I talk about future trends that will influence how our work gets done. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.476	Retain the null hypothesis.
26	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-26. I describe a compelling image of what our future could be like. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.850	Retain the null hypothesis.
27	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-27. I appeal to teachers to share an exciting dream of the future. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.497	Retain the null hypothesis.
28	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-28. I show teachers how their long-term interests can be realized by enlisting in a common vision. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.923	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
29	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-29. I paint the "big picture" of what we aspire to accomplish. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.713	Retain the null hypothesis.
30	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-30. I speak with genuine conviction about the higher meaning and purpose of our work is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.794	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Supplemental table 3. Cross-District Comparison of Individual Teacher LPI Items. Mann-Whitney tests were performed for all 30 individual items from the teacher survey questions across district U1 and district C1. Statistically significant differences ($p<0.05$) were seen for 11 of 30 items (highlighted in yellow).

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-1. I set a personal example of what I expect of students. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.605	Retain the null hypothesis.
2	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-2. I follow through on the promises and commitments that I make. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.198	Retain the null hypothesis.
3	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-3. I ask for feedback on how my actions affect student performance. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.307	Retain the null hypothesis.
4	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-4. I challenge students to try out new and innovative ways to do their work. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.001	Reject the null hypothesis.
5	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-5. I search outside the formal boundaries of my organization for innovative ways to improve what I do. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.011	Reject the null hypothesis.
6	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-6. I ask "what can I learn?" when things don't go as expected. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.507	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
7	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-7. I actively listen to diverse points of view. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.095	Retain the null hypothesis.
8	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-8. I treat students with dignity and respect. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.020	Reject the null hypothesis.
9	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-9. I ensure that students grow in my classroom by learning new skills and developing themselves. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.876	Retain the null hypothesis.
10	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-10. I praise students for a job well done. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.931	Retain the null hypothesis.
11	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-11. I make it a point to let students know about my confidence in their abilities. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.043	Reject the null hypothesis.
12	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-12. I find ways to celebrate accomplishments. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.341	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
13	The distribution of Please rate the extent to which you agree with each of the following statements.-13. I am satisfied with my job. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
14	The distribution of Please rate the extent to which you agree with each of the following statements.-14. I am satisfied with my teaching environment. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
15	The distribution of Please rate the extent to which you agree with each of the following statements.-15. I am satisfied with the current leadership team. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
16	The distribution of Please rate the extent to which you agree with each of the following statements.-16. I am satisfied with the type of leaders we have in our school. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
17	The distribution of Please rate the extent to which you agree with each of the following statements.-17. I make decisions about the selection of other teachers for my school. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.504	Retain the null hypothesis.
18	The distribution of Please rate the extent to which you agree with each of the following statements.-18. I am given the opportunity to teach other teachers. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.140	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
19	The distribution of Please rate the extent to which you agree with each of the following statements.-19. I have an opportunity to teach other teachers about innovative ideas. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.058	Retain the null hypothesis.
20	The distribution of Please rate the extent to which you agree with each of the following statements. -20. Principals, other teachers, and school personnel solicit my advice. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.101	Retain the null hypothesis.
21	The distribution of Please rate the extent to which you agree with each of the following statements.-21. I have the opportunity to collaborate with other teachers in my school. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.502	Retain the null hypothesis.
22	The distribution of Please rate the extent to which you agree with each of the following statements.-22. I work at a school where students come first. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
23	The distribution of Please rate the extent to which you agree with each of the following statements.-23. I have the support and respect of my colleagues. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.023	Reject the null hypothesis.
24	The distribution of Please rate the extent to which you agree with each of the following statements.-24. I have a strong knowledge base in the areas in which I teach. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.831	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
25	The distribution of Please rate the extent to which you agree with each of the following statements.-25. I believe that I am helping students become independent learners. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.987	Retain the null hypothesis.
26	The distribution of Please rate the extent to which you agree with each of the following statements.-26. I perceive that I am making a difference. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.288	Retain the null hypothesis.
27	The distribution of Please rate the extent to which you agree with each of the following statements.-27. I am able to teach as I choose. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
28	The distribution of Please rate the extent to which you agree with each of the following statements.-28. I have the freedom to make decisions on what is taught. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.643	Retain the null hypothesis.
29	The distribution of Please rate the extent to which you agree with each of the following statements.-29. I perceive that I have the opportunity to influence others. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.754	Retain the null hypothesis.
30	The distribution of Please rate the extent to which you agree with each of the following statements.-30. I perceive that I have an impact on other teachers and students. is the same across categories of District.	Independent-Samples Mann-Whitney U Test	.315	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Supplemental Table 4. Cross-National Differences on Individual LPI Items for Teachers. Mann-Whitney tests were performed for all 30 individual items from the teacher survey questions across participant US and Canadian teachers. Statistically significant differences ($p<0.05$) were seen for 11 of 30 items (highlighted in yellow). US teachers had higher scores on 1 LPI question and Canadian teachers having higher scores on 1 LPI item. Canadian teachers had higher scores on all 4 satisfaction questions, and 4 SPES questions.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-1. I set a personal example of what I expect of students. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.841	Retain the null hypothesis.
2	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-2. I follow through on the promises and commitments that I make. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.769	Retain the null hypothesis.
3	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-3. I ask for feedback on how my actions affect student performance. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.384	Retain the null hypothesis.
4	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-4. I challenge students to try out new and innovative ways to do their work. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.010	Reject the null hypothesis.
5	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-5. I search outside the formal boundaries of my organization for innovative ways to improve what I do. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.113	Retain the null hypothesis.
6	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-6. I ask "what can I learn?" when things don't go as expected. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.917	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
7	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-7. I actively listen to diverse points of view. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.862	Retain the null hypothesis.
8	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-8. I treat students with dignity and respect. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.013	Reject the null hypothesis.
9	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-9. I ensure that students grow in my classroom by learning new skills and developing themselves. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.269	Retain the null hypothesis.
10	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-10. I praise students for a job well done. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.768	Retain the null hypothesis.
11	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-11. I make it a point to let students know about my confidence in their abilities. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.063	Retain the null hypothesis.
12	The distribution of Please rate how frequently you engage in the following leadership behaviors. Be as honest and as...-12. I find ways to celebrate accomplishments. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.218	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
13	The distribution of Please rate the extent to which you agree with each of the following statements.-13. I am satisfied with my job. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
14	The distribution of Please rate the extent to which you agree with each of the following statements.-14. I am satisfied with my teaching environment. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
15	The distribution of Please rate the extent to which you agree with each of the following statements.-15. I am satisfied with the current leadership team. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
16	The distribution of Please rate the extent to which you agree with each of the following statements.-16. I am satisfied with the type of leaders we have in our school. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
17	The distribution of Please rate the extent to which you agree with each of the following statements.-17. I make decisions about the selection of other teachers for my school. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.478	Retain the null hypothesis.
18	The distribution of Please rate the extent to which you agree with each of the following statements.-18. I am given the opportunity to teach other teachers. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.062	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
19	The distribution of Please rate the extent to which you agree with each of the following statements.-19. I have an opportunity to teach other teachers about innovative ideas. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.056	Retain the null hypothesis.
20	The distribution of Please rate the extent to which you agree with each of the following statements. -20. Principals, other teachers, and school personnel solicit my advice. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.025	Reject the null hypothesis.
21	The distribution of Please rate the extent to which you agree with each of the following statements.-21. I have the opportunity to collaborate with other teachers in my school. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.648	Retain the null hypothesis.
22	The distribution of Please rate the extent to which you agree with each of the following statements.-22. I work at a school where students come first. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
23	The distribution of Please rate the extent to which you agree with each of the following statements.-23. I have the support and respect of my colleagues. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.174	Retain the null hypothesis.
24	The distribution of Please rate the extent to which you agree with each of the following statements.-24. I have a strong knowledge base in the areas in which I teach. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.424	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
25	The distribution of Please rate the extent to which you agree with each of the following statements.-25. I believe that I am helping students become independent learners. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.095	Retain the null hypothesis.
26	The distribution of Please rate the extent to which you agree with each of the following statements.-26. I perceive that I am making a difference. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.387	Retain the null hypothesis.
27	The distribution of Please rate the extent to which you agree with each of the following statements.-27. I am able to teach as I choose. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.
28	The distribution of Please rate the extent to which you agree with each of the following statements.-28. I have the freedom to make decision on what is taught. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.002	Reject the null hypothesis.
29	The distribution of Please rate the extent to which you agree with each of the following statements.-29. I perceive that I have the opportunity to influence others. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.238	Retain the null hypothesis.
30	The distribution of Please rate the extent to which you agree with each of the following statements.-30. I perceive that I have an impact on other teachers and students. is the same across categories of Country.	Independent-Samples Mann-Whitney U Test	.082	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Supplemental Table 5. Descriptive statistics for demographics across districts U1 and C1. This table lists average ranks for districts U1 and C1 for demographic variables

	District	N	Mean Rank	Sum of Ranks
-Gender-Male	1	181	147.94	26778.00
	5	105	135.84	14263.00
	Total	286		
-Age	1	176	134.81	23727.00
	5	101	146.30	14776.00
	Total	277		
-Years of Teaching / Administrative Experience	1	176	125.21	22037.50
	5	101	163.02	16465.50
	Total	277		
-Education Level	1	176	157.27	27679.50
	5	101	107.16	10823.50
	Total	277		

Supplemental table 6. Test statistics for demographic variables other than ethnicity. Mann-Whitney statistics revealed that years of teaching/administrative experience and education level were significantly different ($p<0.001$) between districts U1 and C1

	-Gender-Male	-Age	-Years of Teaching / Administrative Experience	-Education Level
Mann-Whitney U	8698.000	8151.000	6461.500	5672.500
Wilcoxon W	14263.000	23727.000	22037.500	10823.500
Z	-1.387	-1.158	-3.899	-5.667
Asymp. Sig. (2-tailed)	.166	.247	.000	.000

a. Grouping Variable: District

Supplemental table 7. Assessment of differences in ethnicity across districts U1 and C1. Ranks and Kruskal-Wallis statistics were calculated for ethnicity across districts U1 and C1, this was not significantly different (p=0.13)

Ranks

	District	N	Mean Rank
-Ethnicity	1	181	136.09
	5	100	149.89
	Total	281	

Test Statistics[a,b]

	-Ethnicity
Chi-Square	2.292
Df	1
Asymp. Sig.	.130

a. Kruskal Wallis Test

b. Grouping Variable: District

Supplemental table 8. Tests of significance for demographics. Mann-Whitney statistics revealed that years of teaching/administrative experience and education level were significantly different (p<0.001) between US and Canadian respondents in the 8 participant districts

Test Statistics[a]

	-Gender-Male	-Age	-Years of Teaching / Administrative Experience	-Education Level
Mann-Whitney U	8698.000	8151.000	6461.500	5672.500
Wilcoxon W	14263.000	23727.000	22037.500	10823.500
Z	-1.387	-1.158	-3.899	-5.667
Asymp. Sig. (2-tailed)	.166	.247	.000	.000

a. Grouping Variable: District

Supplemental table 9. Ranks and Kruskal-Wallis statistics were calculated for ethnicity across US and Canadian participants, this was significantly different (p=0.045)

Ranks

	Country	N	Mean Rank
-Ethnicity	US	251	213.02
	Canada	192	233.74
	Total	443	

Test Statistics[a,b]

	-Ethnicity
Chi-Square	4.031
df	1
Asymp. Sig.	.045

a. Kruskal Wallis Test

b. Grouping Variable: Country

Supplemental table 10. MANOVA of cross-district administrator LPI. Analysis showed that after correction for teaching experience and level of education (significant covariates from univariate analysis), there were still no significant differences in administrator LPI across districts U1 and C1

Tests of Between-Subjects Effects – District U1 vs C1, Administrator LPI

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	Model the way P	.096[a]	3	.032	.320	.811
	Challenge the process admin	.467[b]	3	.156	1.210	.321
	Enable others to act admin	.141[c]	3	.047	.407	.749
	Encourage the heart admin	.657[d]	3	.219	.883	.460
	Inspire a shared vision admin	1.503[e]	3	.501	1.897	.149

Supplemental table 11. MANOVA – Cross District teacher LPI & SPES. This analysis corrects for years of experience and level of education covariates demonstrates that while differences remained in teacher's "challenge the process" LPI sub score, satisfaction, and "professional growth" SPES sub score, the overall SPES score and autonomy sub scores were previously shown to be significant were confounded by covariates and lost their significance while "status" SPES sub score was not originally reported as significant but became significant after correction for covariates.

Tests of Between-Subjects Effects

Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Model the way teachers	.558[a]	3	.186	.862	.462
Challenge the process teachers	4.104[b]	3	1.368	3.435	.018
Enable others to act teachers	.159[c]	3	.053	.311	.818
Encourage the heart teachers	.983[d]	3	.328	1.044	.374
Satisfaction teachers	20.141[e]	3	6.714	6.637	.000
Decision-making teachers	2.802[f]	3	.934	.984	.401***
Professional growth teachers	6.452[g]	3	2.151	2.794	.042
Status teachers	3.125[h]	3	1.042	2.943	.034**
Self efficacy teachers	2.541[i]	3	.847	1.597	.191
Autonomy teachers	6.832[j]	3	2.277	1.586	.194***
Impact teachers	2.511[k]	3	.837	1.229	.300
teacherLPI	5.539[l]	3	1.846	1.425	.237
teacherSPES	3.810[m]	3	1.270	1.594	.192***

** became statistically significant after correction for covariates

*** no longer statistically significant after correction for covariates

Supplemental table 12. MANOVA – Cross-national administrator LPI. Analysis showed that after correction for ethnicity, teaching experience and level of education, there were still no significant differences in administrator LPI across 4/5 subscales but that the "challenge the process" sub score which was previously not reported as significant became significant upon correction for covariates.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	Model the way P	.299[a]	4	.075	.478	.752
	Challenge the process admin	2.309[b]	4	.577	2.914	.027**
	Enable others to act admin	.073[c]	4	.018	.161	.958
	Encourage the heart admin	1.160[d]	4	.290	1.074	.376
	Inspire a shared vision admin	3.112[e]	4	.778	1.918	.117

** became statistically significant after correction for covariates

Supplemental table 13. MANOVA – Cross-national teacher LPI & SPES. Analysis showing that after correction for ethnicity, teaching experience and level of education that there are still no significant differences in teacher LPI subscale scores and that differences in satisfaction, overall SPES score and the autonomy subscale were unchanged but that the "decision-making" SPES sub score which was previously reported as significant lost significance upon correction for covariates.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	Model the way teachers	.782[a]	4	.195	.896	.467
	Challenge the process teachers	1.761[b]	4	.440	1.082	.366
	Enable others to act teachers	.313[c]	4	.078	.458	.767
	Encourage the heart teachers	1.810[d]	4	.453	1.502	.202
	Satisfaction teachers	26.338[e]	4	6.584	6.184	.000
	Decisionmaking teachers	4.762[f]	4	1.190	1.201	.311***
	Professional growth teachers	6.718[g]	4	1.679	2.167	.073
	Status teachers	2.819[h]	4	.705	2.035	.090
	Self efficacy teachers	5.001[i]	4	1.250	2.366	.053
	Autonomy teachers	25.241[j]	4	6.310	4.519	.001
	Impact teachers	2.572[k]	4	.643	.897	.466
	teacherLPI	5.923[l]	4	1.481	1.691	.152
	teacherSPES	8.693[m]	4	2.173	3.341	.011

*** no longer statistically significant after correction for covariates

H. Appendices

Appendix A. IRB Approval Letter

DEPAUL
UNIVERSITY

Office of Research Services
Institutional Review Board
1 East Jackson Boulevard
Chicago, Illinois 60604-2201
312-362-7593
Fax: 312-362-7574

Research Involving Human Subjects
NOTICE OF INSTITUTIONAL REVIEW BOARD ACTION

To: Anya Niazov, Graduate Student, College of Education

Date: February 11, 2016

Re: Research Protocol # AN062515EDU-R2

"Leading the Way to Transforming Leadership Across Nations: Assessment of Differences Between Administrator and Teacher Leadership Behaviors in Large Secondary School Districts in the United States and Canada"
Please review the following important information about the review of your proposed research activity.

Review Details
This submission is an Amendment. Amendment R2 involves: 1) a protocol title change; 2) the addition of nine letters of support from various performance/recruitment sites; and 3) revisions to the Exempt Application and the Information Sheet to reflect these changes.

Your research project continues to meet the criteria for Exempt review under 45 CFR 46.101 under the following category:
(2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless: (i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

Approval Details
Your research Amendment was reviewed and approved on February 11, 2016.
Number of approved participants: 2200
You should not exceed this total number of subjects without prospectively submitting an amendment to the IRB requesting an increase in subject number.
Funding Source: None

Approved Performance sites: (Listed in original letter, omitted here for anonymity)

Reminders

Under DePaul's current institutional policy governing human research, research projects that meet the criteria for an exemption determination may receive administrative review by the Office of Research Services Research Protections staff. Once projects are determined to be exempt, the researcher is free to begin the work and is not required to submit an annual update (continuing review). As your project has been determined to be exempt, your primary obligation moving forward is to resubmit your research materials for review and classification/approval when making changes to the research, but before the changes are implemented in the research. **All changes to the research must be reviewed and approved by the IRB or Office of Research Services staff**. Changes requiring approval include, but are not limited to, changes in the design or focus of the research project, revisions to the information sheet for participants, addition of new measures or instruments, increasing the subject number, and any change to the research that might alter the exemption status (either add additional exemption categories or make the research no longer eligible for an exemption determination).

Once the project is complete, you should submit a final closure report to the IRB.

The Office of Research Services would like to thank you for your efforts and cooperation and wishes you the best of luck on your research. If you have any questions, please contact me by telephone at (312) 362-6168 or via email at jbloom8@depaul.edu.

For the Board,

Jessica Bloom, MPH
Research Protections Coordinator
Office of Research Services

Cc: Joseph Gardner, Ph.D., Faculty Sponsor, College of Education

Appendix B: Districts / Boards External Researcher and Ethics Approval

Individual district IRB approval letters were obtained from each of the 8 participating districts: U1, U2, U3, U4, C1, C2, C3, and C4. In pursuance to these approvals, the district identities have been de-identified in this report and therefore the individual district IRB clearances are not reproduced in this section of the appendix to maintain anonymity.

If copies of the individual district approvals are required, please contact me and I will provide copies of these approval letters (after receiving permission from the districts).

Below is a table detailing the items that were included in district applications to obtain approval:

• Proposal (10-30 pages, including research procedures, hypothesis, recruitment strategy, data management and disposal plan, literature review)
• Copy of DePaul IRB approval/application +/- district's internal district ethics review process application
• CITI training completion & certification / NIH proof of training & certification
• District specific information letters and consents
• Adult participant voluntary invitation letters
• Table of time requirement and estimated participant numbers
• List of preferred schools (meeting my inclusion criteria – each school within each district individually researched & determined)
• Recommendation letters / letters of support (including mandatory support letters from each Chief within each district)
• Survey instruments (needed to be modified per each district rules)
• Benefit to participants & district & alignment with district's strategic multi-year plan and/or research priorities (modified for each district)
• References

Appendix C: Demographic Survey

1. Gender: Female / Male

2. Ethnicity (US Survey): African American / Native American / Hispanic / Asian/Pacific Islander / Caucasian / Multi-Racial / Other

3. Ethnicity (Canadian Survey): African Canadian / Aboriginal / Latin / Asian /Pacific Islander / Caucasian / Multi-Racial / Other

4. Age: 24 or younger / 25-29 / 30-34 / 34-39 / 40-44 / 45-49/ 50 or older

5. Education Level: Associates / Undergraduate / Masters / Doctorate

6. Years of (teaching / administrative) experience: 0-5 / 6-10 / 11-15 / 16-20 / 21 or more

7. Select your school: (drop down menu of schools with associated de-identified code that the researcher was not be privy to)

For those who did not want to specify their school, there was be an options of "Prefer not to disclose" and "Other" for those not currently affiliated with any of the approved schools. Because participants were directed to the survey link specific to their district, only schools approved by the district were be shown

Appendix D: Administrator Survey

Please rate how frequently you engage in the following leadership behaviors. Be as honest and as accurate as you can be. Please answer in terms of how you typically behave on most days, on most projects, and with most people.

Rate each statement on the following scale:

1=Almost Never 2=Seldom 3=Occasionally 4=Usually 5=Almost Always

1. I set a personal example of what I expect of teachers.
2. I follow through on the promises and commitments that I make.
3. I ask for feedback on how my actions affect teachers' performance.
4. I spend time and energy making certain that the teachers I work with adhere to the principles and standards we have agreed on
5. I build consensus around a common set of values for running our organization.
6. I am clear about my philosophy of leadership.
7. I challenge teachers to try out new and innovative ways to do their work.
8. I search outside the formal boundaries of my organization for innovative ways to improve what we do.
9. I ask "what can we learn?" when things don't go as expected.
10. I seek out challenging opportunities that test my own skills and abilities.
11. I make certain that we set achievable goals, make concrete plans, and establish measurable milestones for the projects and programs that we work on.
12. I experiment and take risks, even when there is a chance of failure.
13. I actively listen to diverse points of view.
14. I treat teachers with dignity and respect.
15. I ensure that teachers grow in their job by learning new skills and developing themselves.
16. I develop cooperative relationships among the teachers I work with.
17. I support the decisions that teachers make on their own.
18. I give teachers a great deal of freedom and choice in deciding how to do their work.
19. I praise teachers for a job well done.
20. I make it a point to let teachers know about my confidence in their abilities.
21. I find ways to celebrate accomplishments.
22. I make sure that teachers are creatively rewarded for their contributions to the success of our projects.
23. I publicly recognize teachers who exemplify commitment to shared value.
24. I give the members of the team lots of appreciation and support for their contributions.
25. I talk about future trends that will influence how our work gets done.
26. I describe a compelling image of what our future could be like.
27. I appeal to teachers to share an exciting dream of the future.
28. I show teachers how their long-term interests can be realized by enlisting in a common vision.
29. I paint the "big picture" of what we aspire to accomplish.
30. I speak with genuine conviction about the higher meaning and purpose of our work

Appendix E: Teacher survey

Please rate how frequently you engage in the following leadership behaviors. Be as honest and as accurate as you can be. Please answer in terms of how you typically behave on most days, on most projects, and with most people.

Rate each statement on the following scale:

1=Almost Never 2=Seldom 3=Occasionally 4=Usually 5=Almost Always

1. I set a personal example of what I expect of students.
2. I follow through on the promises and commitments that I make.
3. I ask for feedback on how my actions affect student performance.
4. I challenge students to try out new and innovative ways to do their work.
5. I search outside the formal boundaries of my organization for innovative ways to improve what I do.
6. I ask "what can I learn?" when things don't go as expected.
7. I actively listen to diverse points of view.
8. I treat students with dignity and respect.
9. I ensure that students grow in my classroom by learning new skills and developing themselves.
10. I praise students for a job well done.
11. I make it a point to let students know about my confidence in their abilities.
12. I find ways to celebrate accomplishments.

Please rate the extent to which you agree with each of the following statements.

Rate each statement on the following scale:

1=Strongly Disagree 2=Disagree 3=Neutral 4=Agree 5=Strongly Agree

13. I am satisfied with my job.
14. I am satisfied with my teaching environment.
15. I am satisfied with the current leadership team.
16. I am satisfied with the type of leaders we have in our school.
17. I make decisions about the selection of other teachers for my school.
18. I am given the opportunity to teach other teachers.
19. I have an opportunity to teach other teachers about innovative ideas.
20. Principals, other teachers, and school personnel solicit my advice.
21. I have the opportunity to collaborate with other teachers in my school.
22. I work at a school where students come first.
23. I have the support and respect of my colleagues.
24. I have a strong knowledge base in the areas in which I teach.
25. I believe that I am helping students become independent learners.
26. I perceive that I am making a difference.
27. I am able to teach as I choose.
28. I have the freedom to make decision on what is taught.
29. I perceive that I have the opportunity to influence others.
30. I perceive that I have an impact on other teachers and students.

Appendix F: Participant Information Letter

INFORMATION SHEET FOR PARTICIPATION IN RESEARCH STUDY

Leading the Way to Transforming Leadership Across Nations:

Assessment of differences between administrator and teacher leadership behaviors in large secondary school districts, in the United States and Canada

Principal Investigator: Anya Niazov, Doctoral Candidate

Department of Educational Leadership, College of Education

Institution: DePaul University, USA

Faculty Advisor: Dr. Joseph Gardner, Program Director

Department of Curriculum Studies, College of Education

We are conducting a research study because we are trying to learn more about the differences between administrators' and teachers' leadership beliefs and behaviors in large districts in the U.S. and in Canada. In order to address this issue, previously-validated surveys will be administered to better understand (1) leadership practices and (2) teacher empowerment among secondary school teachers and administrators from the (district) and (other district(s)). We are asking you to be in the research because you are a full-time licensed teacher, principal, or vice principal in a secondary school in either one of the districts. You must be age 21 or older to be in this study. This study is not approved for the enrollment of people under the age of 21. Additionally, you must read and understand English and have access to a computer and the internet. If you agree to be in this study, you will be asked to fill out a short online survey. The survey will include questions about your beliefs about your leadership and empowerment behaviors and practices. You will also be asked demographic information such as age, gender, years of experience and school name.

This study will take about 5-10 minutes of your time and consist of 30 statements about leadership that that you will evaluate on a 5-point Likert scale. Research data collected from you will be anonymous and will not be linked to you. Please do this on your own spare time.

Your individual participation is voluntary, which means that your consent is being sought to collect information in a survey format. There will be no negative consequences if you decide not to participate or change your mind later after you begin the study. You can withdraw your participation at any time prior to submitting your survey. Your decision whether or not to be in the research will not affect your employment at your school. If you change your mind later while answering the survey, you may simply exit the survey. Once you submit your responses, we will be unable to remove your data later from the study because all data is anonymous and we will not know which data belongs to you.

No personally identifiable data will be used in this research. All data will be de-identified and coded based on (1) country (2) school and (3) whether the participant is a teacher or administrator, there will be no way to identify individual participants' responses. Additionally, the IP address tracking feature on the survey software is disabled so that individual participants' identities/locations will not be known. All data will be kept only in coded aggregate form with no identifiable information accessible to the researcher. The raw (individual level) data will be housed at DePaul University (password protected and encrypted so that only I will have access to the data) and will be destroyed after completion of the study.

Please Note: This study has been approved by DePaul University Institutional Review Board and (district IRB) and relevant school principals

If you have questions, concerns, or complaints about this study or you want to get additional information or provide input about this research, please contact Anya Niazov via phone or email or my faculty adviser Joseph Gardner.

If you have questions about your rights as a research subject you may contact Susan Loess-Perez, DePaul University's Director of Research Compliance, in the Office of Research Services at 312-362-7593 or by email at sloesspe@depaul.edu. You may also contact DePaul's Office of Research Services if a) your questions, concerns, or complaints are not being answered by the research team, b) you cannot reach the research team, or c) you want to talk to someone besides the research team.

You may print this information for your records.

By proceeding with the survey, I acknowledge that I have read and familiarized myself with the information regarding the research study and that I voluntary agree to participate in this research.

☐ I have read the above information and agree to participate in the research study.

☐ I do not consent to participate in the study.

Note: No name or personally identifiable information is collected as part of this study

Appendix G: Principal school recruitment letter

Dear Principal,

Effective leadership behaviors in the classroom and from school administrators are essential to contributing to student success. My doctoral research seeks to gain new knowledge about administrators' and teachers' leadership beliefs and behaviors by studying this critical but often under-appreciated aspect in secondary schools from large urban districts. I will specifically address teacher and administrator attitudes toward leadership, leadership behaviors, and their relationship with teacher empowerment using previously validated research surveys (Leadership Practices Inventory, LPI and School Participant Empowerment Scale, SPES) administered to teachers and administrators at multiple schools in the participant districts. The data from these surveys will be used to begin to understand (using quantitative techniques) how leadership differs amongst teachers and administrators within and between districts, and how administrative leadership impacts teacher leadership and teacher empowerment. This is important because these leadership behaviors have been shown to positively correlate with a number of academic outcomes and to date there has not been a study that addresses these differences cross-nationally.

Secondary schools from each district have been chosen based on a number of metrics and your school was deemed appropriate for participation in this study. Within your school, participant study inclusion criteria include that individuals must be full-time employees who work in your public high school. All participants must be 21 years or older, must be able to read and understand English, and have access to a computer and the internet. Eligible participants will be either (1) licensed teachers or (2) administrators. Administrators include both principals and vice principals.

The online surveys employed should take only 5-10 minutes for teachers and administrators to complete. All data will be collected in a de-identified manner such that individuals will not be able to be connected to their survey responses and study reports will talk about aggregate district results rather than school level results. All data collected will be used only to answer specific research questions in the context of a doctoral dissertation and will not be shared with any other researchers or institutions.

After the study is completed, schools that participate in the study will receive both a copy of the final report describing the relationship between teacher and administrator leadership within the (district) schools and cross-nationally as well as a report specifically addressing the scores of teachers and administrators at the individual school. These data should provide important insights into the types of leadership that are present at your school, how your school's administrative leadership behaviors impact teachers, and give you benchmark data against other schools in the (districts) schools which you can use to create actionable interventions to improve your school's leadership culture.

I hope that you will see the value in this important work and look forward to working with you and your staff to better understand and improve school leadership

Sincerely,

Anya Niazov

Doctoral Candidate, DePaul University

Appendix H: Sample recruitment flyer/email

BE A RESEARCH PARTICIPANT IN THE (DISTRICT) **SCHOOL LEADERSHIP SURVEY**!

Be a part of this exciting new research!

The (district) IRB has approved a new cross-national study of teacher and administrator leadership and **your school has been selected** as one of the participant sites.

This *study will address leadership* behaviors in the classroom and from school administrators. To gain deeper insights into this important area of study, my doctoral research seeks to gain new knowledge about the differences among administrators' and teachers' beliefs about leadership behaviors and empowerment. *Your participation is voluntary* but will help make the research meaningful and provide important information to ***improve your school and district.***

My name is Anya Niazov and I am a doctoral candidate in the department of Educational Leadership, College of Education at DePaul University in Chicago, Illinois. I am asking you to ***take 5 minutes to complete this online survey*** *which will help your school and district learn more about effective leadership techniques, how administrator leadership affects teachers, and how leadership behaviors differ cross-nationally.*

The survey link is:

Website link was inserted here

This survey is open **from (1 month period of districts' study approval)**.
Please log on now to participate.

Thank you,
Anya Niazov,
DePaul University, Chicago, IL

REFERENCES

Acker-Hocevar, M. & Touchton, D. (1999). A model of power as social relationships: Teacher leaders describe the phenomena of effective agency in practice. (Report No.SP 040 191). Washington, DC: National Center for Research on Teaching and Teacher Education (ERIC No. ED 456 108)

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckman (Eds.), *Action-control: From cognition to behavior* (pp. 11-39). Heidelberg: Springer.

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The Handbook Of Attitudes* (pp. 173-221). Mahwah, NJ: Erlbaum.

Ashton, P. T., & Webb, R. B. (1986). *Making a difference: Teachers' sense of efficacy and Student achievement*. New York: Longman

Baker K. (2007). Are International Tests Worth Anything? *Phi Delta Kappan*, 89(2), 101-104.

Baker, D. & LeTendre, G. (2005). National Differences, Global Similarities: World Culture and the Future of Schooling. California: Stanford University Press

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191–215.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Clinical and Social Psychology*, 4, 359-373.

Bandura, A. (1989). Regulation of cognitive processes through perceived self efficacy. *Developmental Psychology*, *25*(5), 725-729.

Barnett, H., Craven, R. & Marsh, A. (2005). What type of school leadership satisfies teachers? A mixed method approach to teachers' perceptions of satisfaction, *Australian Association for Research in Education.*

Barnett, K., McCormick, J., & Conners, R. (2003). Vision, Relationships and Teacher Motivation: A Case Study. *Journal of Educational Administration,* 41(1), 55-73.

Barth, R. (2007). The teacher leader. In H. Ackerman & S. Mackenzie (Eds.*), Uncovering teacher leadership: Essays and voices from the field*. (pp. 9-36). Thousand Oaks: Corwin Press

Barth, R. S. (2001). *Learning by heart*. San Francisco, CA: Jossey-Bass.

Bason, M.R., & Frase, L. (2004). Creating optimal work environments: exploring teacher flow experiences. *Mentoring and Tutoring*, 12 (2).

Bass, B. & Avolio, B. (1994). Introduction. In B. Bass and B. Avolio (Eds.), *Improving organizational effectiveness through transformational leadership* (pp. 1-9). Thousand Oaks, CA: Sage.

Bass, B. (1985). *Leadership and performance beyond expectations*. New York: Free Press.

Berliner, D.C. & Biddle, B.J. (1996). *The manustructured crisis: Myth, fraud, and the attack on public schools.* New York: Perseus

Berry, B., Daughtrey, A., & Wieder, A. (2010). Teacher leadership: Leading the way to effective teaching and learning. Retrieved from Center for Teaching Quality: http://www.teachingquality.org

Blase, J., & Blase, J. (1999). Principals' instructional leadership and teacher development: teachers' perspectives. *Educational Administration Quarterly,* 35(3), 349-78.

Blase, J., & Kirby, P. C. (2008). *Bringing out the best in teachers: What effective principals do*. Thousand Oaks, CA: Corwin Press

Boe, E. & Shin, S. (2005). Is the United States Really Losing the International Horse Race in Academic Achievement? *Phi Delta Kappan,* 86(9), 88-695.

Bogler, R (2001). The Influence of Leadership Style on Teacher Job Satisfaction Educational *Administration Quarterly,* 37(5): 662-683

Bousquet, S. (2012). Teacher burnout: causes, cures, and prevention. ERIC online submission No. ED534527

Bracey, G. (2009). Big tests: What ends do they serve? *Educational Leadership,* 67 (3), 32 – 37.

Bray, M. & Thomas, R. M. (1995). Levels of comparison in educational studies: Different insights from different literatures and the value of multilevel analyses. *Harvard Education Review*, 65(3), 472-490

Bray, M., Adamson, B., & Mason, M. (2007). Comparative education research: Approaches and methods. Hong Kong: Comparative Education Research Centre, The University of Hong Kong.

Brooks, S., Giles, C., Jacobson, S., C., Johnson, L., & Yimaki, R. (2007). Successful leadership in three high-poverty urban elementary schools, *Leadership and Policy in Schools,* 6(4), 291-317.

Bryant & Yan (2010) Urban School Teachers' Self-Efficacy Beliefs and AYP Status, *The Charter Schools Resource Journal, v6*

Burns, J. M. (1978). *Leadership.* New York: Harper and Row.

Bushaw, S. & Lopez, J. (2010). A Time for Change: The 42nd Annual Phi Delta Kappa/ Gallup Poll of the Public's Attitudes toward the Public Schools, *Phi Delta Kappan, 92(1), 9-26*

Canadian Education Association. (2009). Learning about leadership from school turnaround efforts in Ontario. Ontario, Toronto: Leithwood, K., & Strauss, T.

Cavanagh, S. & Manzo, K.K. (2009). International Exams Yield Less-Than-Clear Lessons. *Education Week,* 28(29), 1-17.

Cavanagh, S. (2012). Complex Policy Options Abound Amid International Comparisons. *Education Week,* 31(16), 6-10.

Charles, & Mertler. (2002*). Introduction to educational research* (4th ed.). Boston: Allyn and Bacon.

Chin, J. (2007). Meta-analysis of transformational school leadership effects on school outcomes in Taiwan and the USA. *Asia Pacific Education Review*, *8*(2), 166-177.

Choy, S. P., Bobbitt, S. A., Henke, R. R., Medrich, E. A., Horn, L J., and Lieberman, J. (1993). *America's Teachers: Profile of a Profession*. Washington, DC.

Clandinin, DJ, Schaefer L, Long JS, Steeves P, McKenzie-Robblee S, Pinnegar E, Wnuk S, and Downey CA (2013). Teacher Education: A Question of Sustaining Teachers. p. 251-262. *Preparing Teachers for the 21st Century*. Eds. Xhu, Xiodong & Ziechner, Kenneth. Springer: Berlin.

Coad, A. & Berry, A. (1998). Transformational leadership and learning orientation. *Leadership & Organization Development Journal,* 19(3), 164-172.

Cochran-Smith, M., & Lytle, S. L. (2009*). Inquiry as stance: Practitioner research for the next generation*. New York: Teachers College Press.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles: Sage.

Cross, C. T. (2010). *Political education: National policy comes of age*. New York: Teachers College Press.

Crossley, M. (2012). Comparative education and research capacity building: Reflections on international transfer and the significance of context. *Journal of International and Comparative Education*, 1(1), 4-12

Datnow, A. (2008). The co-construction of educational reform: The intersection of federal, state, and local contexts. In M. Shinn, H. Yoshikawa, M. Shinn, H. Yoshikawa (Eds.) , *Toward positive youth development: Transforming schools and community programs* (pp. 271-291). New York, NY, US: Oxford University Press.

Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future.* New York: Teachers College Press

Darling-Hammond, L. (2010-2011). Soaring systems: High flyers all have equitable funding, shared curriculum, and quality teaching. *American Educator, 34(4), 20-23, 53.*

Darling-Hammond, L. (2011). Soaring Systems: High Flyers All Have Equitable Funding, Shared Curriculum, and Quality Teaching. *American Educator*, 34(4), 20-23, 53.

Davidson, Betty M.; Dell, Geralyn L (2003). A School Restructuring Model: A Tool Kit for Building Teacher Leadership. Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL)

Day, C. (2013). *Teacher Quality in the Twenty First Century: New Lives, Old Truths. Preparing Teachers for the 21st Century*. Eds. Xhu, Xiodong & Ziechner, Kenneth. Springer: Berlin. 21-39.

Dillman, D., & Schaefer, D. (1998). Development of a standard e-mail methodology: Results of an experiment. *Public Opinion Quarterly, 62*, 378-397. Retrieved from http://www/sesrc.wsu.edu/dillman/

DiPaola, M.F. & Hoy, W.K. (2008). *Principals improving instruction: Supervision, evaluation, and professional development*. Boston: Pearson/Allyn & Bacon.

Document Reproduction Service ED 352710

Dolton, P & Marcenaro-Gutierrez, O. (2013). Varkey GEMS Foundation Global Teacher Status Index.

Dong, N., & Cravens, X. C. (2011). Identifying malleable school-level learning conditions for mathematics achievement from TIMSS 1999 and 2007. Paper presented at the Annual Convention of American Educational Research Association, New Orleans, LA.

Duncan D. Nulty. (2008). The adequacy of response rates to online and paper surveys: what can be done? *Assessment & Evaluation in Higher Education* 33(3), 301–314.

Dunne, T., & Fee, K. (2008). Differences in Educational Attainment across States. Economic Trends (07482922), 20.

Edington E.D., Benedetto, D, & Raymond, R. (1988). Principal Leadership Style and Student Achievement in Small and Rural Schools of New Mexico. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA. Abstract retrieved from http://eric.ed.gov/?id=ED295770

Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership, 37*(1), 15-24.

Education Reporter. (2011). Why National Curriculum is Bad for America: A Critical Response to the Shanker Institute Manifesto and the U.S. Department of Education's Initiative to Develop National Curriculum and Assessments Based on National Standards (N 305). Article posted to www.eagleforum.org/educate/2011/june11/focus.html

Fensterwald, J. (2013). California looks to Ontario schools' reformer for guidance. EdSource: Highlighting strategies for student success. Retrieved from http://edsource.org/2013/california-looks-to-ontario-schools-reformer-for-guidance/30644#.VWavDc9Viko

Ferrari, G.R.F. (ed). (2000). *Plato: The Republic*, T. Griffith (trans.). Cambridge: Cambridge University Press.

Fink, E. & Resnick, (2001). Developing principals as instructional leaders.

Finn. C., E. Jr. (1997) "Paternalism Goes to School" in The New Paternalism: Supervisory for the United States. Organization for Economic Cooperation and Development.

Fowler, F. C. (2013) *Policy studies for educational leaders: an introduction* Boston: Allyn & Bacon/Pearson.

Fullan, M. (2001). Implementing change at the building level. *Critical and Emerging Issues in Education Leadership.* August edition, University of Toronto.

Fullan, M. (2007). *Turnaround leadership*. San Francisco: Jossey-Bass.

Fullan, M. (2008). *The six secrets of change: What the best leaders do to help their organizations survive and thrive*. San Francisco: Jossey-Bass.

Gaffield, C. (1994). Children's Lives and Academic Achievement in Canada and the United States. *Comparative Education Review, 38(1), 36-64.*

Gall, J. P., Gall, M. D., & Borg, W. R. (2010). *Applying educational research: How to read do, and use research to solve problems in practice* (6th ed). New York: Allyn & Bacon

Garity, Thomas (2011). Doctoral Dissertation Lincoln Memorial University. Correlation between transformational leadership practices of teachers and administrators of a large suburban newly established high school

Gerstner, C. & Day, D. (1994). Cross-cultural Comparison of Leadership Prototypes. *Leadership quarterly, 5, 121-34.*

Ghaemi,F & Sabokrouh, F (2014). Teacher Empowerment And Its Relationship To Job Satisfaction: A Case Study In Mazandaran University. *International Journal Of Language Learning And Applied Linguistics World* 5(2), 287-298

Gibson, S., & Dembo, M.H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76(4),* 569-582.

Glickman, C. D. (2002). *Leadership for learning: How to help teachers succeed*. Alexandria, VA: ASCD.

Goddard, J., Hoy, H., & Woolfolk, A. (2000). Collective teacher efficacy: its meaning, measure, and impact on student achievement. *American Education Research Journal,* 37(2), 479-507.

Goddard, R. D., Hoy, W. K., & Woolfolk Hoy, A. (2004). Collective efficacy: Theoretical developments, empirical evidence, and future directions. *Educational Researcher,* 33(3), 3-13.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). Highlights From TIMSS 2007: Mathematics and Science Achievement of U.S. Fourth and Eighth-Grade Students in an International Context. (NCES 2009-1001). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Gronn, P. (2002). Distributed leadership. In K. Leithwood & P. Hallinger (Eds.*), International handbook of educational leadership & administration* 653-696. Boston: Kluwer

Gutek. G.L. (1993). *American Education in a Global Society: Internationalizing Teacher Education*. Chicago, IL: Waveland Press, Inc.

Hallinger, P., & Heck, R. (1998). Exploring the principal's contribution to school effectiveness: 1980 – 1999. *School Effectiveness and School Improvement*, 9(2), 157-191.

Hallinger, P.J. (1983).. Assessing the instructional management b behavior of principals. Doctoral dissertation Retrieved from http://search.proquest.com/docview/303278128?accountid=2837

Halverson, R, Kelley C, & Shaw J (2014). "A CALL for improved school leadership." *Phi Delta Kappan* 95.6: 57-60.

Hargreaves, A. & Shirley, D. (2012). *The Global Fourth Way: The Quest for Educational Excellence*. Thousand Oaks, CA: Corwin

Hargreaves, A., & Shirley, D. (2009). The Persistence of Presentism. *Teachers College Record, 111*, 11, 2505-2534.

Hatch, T., White, E.M. & Faigenbaum, D. (2005). Expertise, Credibility, and Influence: How Teachers can Influence Policy, Advance Research, and Improve Performance. *Teachers College Record, 107(5), 1004-1035*

Heck, R. H. (2000). Examining the impact of school quality on school outcomes and improvement: A value-added approach. *Educational Administration Quarterly, 36(4), 513-552*

Heck, R., & Hallinger, P. (2009). Assessing the Contribution of Distributed Leadership to School Improvement and Growth in Math Achievement. *American Educational Research Journal, 46(3),* 659-689.

Hipp, K. & Bredeson, P. (1995). Exploring connections between teacher efficacy and principals' leadership behavior. *Journal of School Leadership,* 5(2), 136-150.

Hipp, K. (1996). Teacher efficacy: Influence of Principal Leadership Behavior. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Hipp, K, & Bredeson, P. (1995). Exploring connections between teacher efficacy and principals' leadership behaviors. *Journal of School Leadership,* (5):136–150.

Hofstede, G. H. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage Publications.

Hofstede, G. H. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. Thousand Oaks, CA: Sage Publications.

House, R. J., Wright, N. S., & Aditya, R. N. (1997). Cross-cultural research on organizational leadership: A critical analysis and a proposed theory. In P. C. Earley & M. Erez (Eds.),

New perspectives in international and industrial organizational psychology (pp. 535–625). San Francisco, CA: New Lexington.

House, R.J., Hanges, P.J., Ruiz -Quintanilla, A. S., Dorfman, P.W., Javidan, M., Dickson, M. W., et al. (1999) "Cultural Influences in Leadership and Organizations: Project GLOBE." In W.H. Mobley, M.J. Gessner and V. Arnold (Eds.*), Advances in Global Leadership*, Vol. 1, pp.171-233. Stamford, Connecticut: JAI Press.

Hulpia, H., Devos, G., & Rosseel, Y. (2009). Development and Validation of Scores on the Distributed Leadership Inventory. *Educational and Psychological Measurement, 69,* 6,1013-1034.

Ingersoll, R. (2007). Short on power, long on responsibility. *Education Leadership,* 65(1), 20-25.

Jackson, T., Burns, J., Bassett, K., & Roberts, R. D. (2010). Teacher leadership: An assessment framework for an emerging area of professional practice. Princeton, NJ: Educational Testing Service.

Jacoby, B. (2004). Advancing education for civic engagement and leadership in a large research university: Excerpts from the report of the Team on Civic Engagement and Leadership at the University of Maryland, College Park. Paper presented at the Association of American Colleges and Universities' Pedagogies of Engagement Conference, Washington, DC.

Jenkins, K. (1988). Metaphor and mindset for educational leadership. *The Educational Forum* 52(2), 143-151.

Kell, M. & Kell, P. (2010). International Testing: Measuring Global Standards or Reinforcing Inequalities. *International Journal of Learning*, 17(9), 485-501.

Kellaghan, T. (2006). What monitoring mechanisms can be used for cross-national (and national) studies? pp 51-63. Cross-national studies of the quality of education: planning their design and managing their impact. Eds Ross, K, and Genovois J. International Institute for Educational Planning: Paris.

Kinder Institute for Urban Resarch. (2012). Houston Education Survey: Public Perceptions in a Critical Time. Houston, TX: Klineberg, S.L, Wu, J., & Douds, K.

Kliebard, H. (2004) *The Struggle for the American Curriculum: 1983 – 1958*. New York: Routledge/Falmer

Kouzes, J. M. & Posner, B. Z. (1997). *The leadership challenge: How to keep getting extraordinary things done in organizations.* San Francisco, CA: Jossey-Bass Publishers.

Kouzes, J. M., & Posner, B. Z. (1995*). The leadership challenge: How to keep getting extraordinary things done in organizations*. San Francisco: Jossey-Bass.

Kouzes, J. M., & Posner, B. Z. (2002). *The leadership challenge* (3rd ed.). San Francisco: John Wiley & Sons.

Kouzes, J. M., & Posner, B. Z. (2002). The leadership practices inventory: Theory and evidence behind the five practices of exemplary leaders (Appendix0512BP ed.). Retrieved from http://media.wiley.com/assets/463/74/ic_jb_appendix.pdf

Kouzes, J. M., & Posner, B. Z. (2007). *The leadership challenge* (4th ed.). San Francisco, CA: Jossey-Bass.

Kouzes, J. M., & Posner, B.Z. (1995). *The leadership challenge: How to keep getting extraordinary things done in organizations.* California: Jossey-Bass.

Kubow, P.K. & Fossum P.R. (2003). *Comparative Education: Exploring Issues in International Context*. Upper Saddle River, NJ: Merrill Prentice Hall.

Leech, D. W., Smith, R., Green, R., & Fulton, C. R. (2003). Exploring Teacher Perceptions of the Leadership Practices of Middle and High School Principals.

Leech D. & Fulton C. (2008) . Faculty Perceptions of Shared Decision Making and the Principal's Leadership Behaviors in Secondary Schools in a Large Urban District. *Education*, v128 n4 p630-644

Leithwood, K. & Jantzi, D. (2000). Principal and teacher leadership effect: a replication. *School Leadership and Management*, 20(4), 415-434.

Leithwood, K. & Jantzi, D. (2006). Transformational school leadership for large-scale reform: effects on students, teachers, and their classroom practices. *School Effectiveness and School Improvement, 17(2),* 201-227.

Leithwood, K. (1994). Leadership for school restructuring. *Educational Administration Quarterly, 30, 498-518.*

Leithwood, K., & Jantzi, D. (2005). *Transformational leadership*. In B. Davies (Ed.), The essentials of school leadership. Thousand Oaks, CA: Sage Publications.

Leithwood, K., & Riehl, C. (2005). What we know about successful school leadership. In W. Firestone & C. Riehl (Eds.), *A new agenda: Directions for research on educational leadership* (pp.22-47). New York, NY: Teachers College Press.

Leithwood, K., Day, C., Sammons, P., Harris, A., & Hopkins, D. (2006). *Successful school leadership: What it is and how it influences pupil learning.*

Leithwood, K., Jantzi, D., & Fernandez, A. (1993). Secondary school teachers' commitment to change: The contributions of transformational leadership. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Leithwood, K., Mascall, B., Strauss,T., Sacks, R., Memon, N. & Yashkina, A. (2007). *Distributing leadership to make schools smarter*. University of Toronto, OISE.

Levin, B. (2008). *How to change 5,000 schools: A practical and positive approach for leading change at every level.* Cambridge, MA: Harvard University Press.

Levin, B. (2011). Comparing Canada and the U.S. on Education. *Education Week.* Retrieved from http://blogs.edweek.org/edweek/futures_of_reform/2011/04/comparing_canada_and_the_us_on_education.html

Levin, B. et Segedin, L. (2011). International Approaches to Secondary Education. Toronto: Higher Education Quality Council of Ontario.

Lightfoot S. (1986). On Goodness of schools: Themes of empowerment. *Peabody Journal of Education*, 63 (3) (1986), pp. 9–28

Lin, E. & Shi, Q. (2014). Exploring Individual and School-Related Factors and Environmental Literacy: Comparing U.S. and Canada Using PISA 2006. *International Journal of Science and Mathematics Education*, 12(1), 73-97.

Litz, D (2011) Globalization and the Changing Face of Educational Leadership: Current Trends and Emerging Dilemmas.

Lorenc, B., Loosveldt, G., Mulry, M. H., & Wrighte D. (2013). Understanding and Improving the External Survey Environment of Official Statistics. Survey Methods: Insights from the Field. Retrieved from http://surveyinsights.org/?p=161

Maden, M. (Ed.). (2001). *Success against the odds, five years on: Revisiting effective schools in disadvantaged areas*. London: Routledge Falmer.

Maeroff, G. I. (1988). *The empowerment of teacher: Overcoming the crisis of confidence* . New York, NY: Teachers College Press.

Marks, Helen M. & Printy, Susan, M. (2003). Principal leadership and school performnace: An integration of transformational and instructional leadership. *Educational Administration Quarterly, 39, 370-397.*

Marzano, R., Waters, T. & McNulty, B. (2005). School Leadership That Works: From Research to Results. Alexander, Virginia: ASCD. Aurora, Colorado: McREL.

Marzano, R., Waters, T., & McNulty, B. (2005). School leadership that works: From research to results. Alexandria, VA: Association of Curriculum and Development

McAdams, R. (1993). *Lessons from Abroad: How Other Countries Educate Their Children*. Lancaster, PA: Technomic Publishing Co.

McMillan, J. H., & Schumacher, S. (2001). Research in education: A conceptual introduction (5th ed.). New York: Longman. Meeting of the Mid-South Educational Research Association in Knoxville. TN.

Merry, J.J. (2013). The Mediocre Performance of U.S. Students on International Education Tests: Are Schools to Blame? (Master's thesis). Retrieved from OhioLINK. (Acession osu1330109955)

Milke, M. (2010). School Choice in Canada: Lessons for America. Heritage Foundation, http://www.heritage.org/research/reports/2010/11/school-choice-in-canada-lessons-for-america

Moller, G., Childs-Bowen, D., & Scrivner, J. (2001). Teachers of the year speak out: Tapping into teacher leadership. Greensboro, NC: South Eastern Regional Vision for Education

Muijs, D., & Harris, A. (2006). Teacher led school improvement: Teacher leadership in the UK. *Teaching and Teacher Education, 22*(8), 961-972.

Mullis, I. V. S., Owen, E.H., & Phillips, G.W. (1990). *America's Challenge: Accelerating Academic Achievement*. Princeton, NJ: Educational Testing Service.

Murray, T.R. (2003). *Blending Qualitative and Quantitative Research Methods in Theses and Dissertations*. Thousand Oaks, CA: Corwin Press

National Association of Elementary School Principals. Rethinking Principal Evaluation: A New Paradigm for Informed Policy and Improved Practice: Executive sumary. 2016. pp 1-8. http://www.naesp.org/sites/default/files/PrincipalEvaluationReportExecutiveSummary.pdf

National Center for Education Statistics: Institute of Education Sciences (1995). TIMSS 1995 Results [Data file]. Retrieved from http://nces.ed.gov/TIMSS/results95.asp

National Commission on Excellence in Education. (1983). *A Nation at Risk. The Imperative for Education Reform.* U.S. Government Printing Office

Nir, A., & Kranot, N. (2006). School principal's leadership style and teachers' self-efficacy. *Planning and Changing, 37,* 205-218.

Noonan, B., & Walker, K. (2008). Trust in the contemporary principalship. *Canadian Journal of Education Administration and Policy*, 85.

O'Donnell, R. J., & White, G. P. (2005). Within the accountability era: Principals' instructional leadership behaviors and student achievement. *NASSP Bulletin,89*(645), 56-71.

Oplatka, I., & Addi-Raccah, A. (2009). Is 'educational leadership' a national-contextual field of study? Some insights from an analysis of the field's major journals. *Educational Leadership: Global contexts and international comparison* (399-418). London: Emerald Publishing Group.

Organization for Economic Co-operation and Development. (2007). Education at a Glance 2007: Highlights. Retrieved from www.oecd.org/dataoecd/4/55/39313286.pdf

Organization for Economic Co-operation and Development. (2010). Strong Performers and Successful Reformers in Education: Lessons from PISA for the United States. Retrieved from http://dx.doi.org/10.1787/9789264096660-en

Organization for Economic Co-operation and Development. (2011). Lessons from PISA for the United States: Strong Performers and Successful Reformers in Education. Retrieved from http://dx.doi.org/10.1787/9789264096660-en

Organization for Economic Co-operation and Development. (2012). Education at a Glance 2012: Highlights. Retrieved from http://dx.doi.org/10.1787/eag_highlights-2012-en

Organization for Economic Co-operation and Development. (2013). Results in Focus 201. Retrieved from http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf

Paine, S. & Schleicher, A. (2011). *What the U.S. can learn from the world's most successful education reform efforts.* New York, NY: McGraw-Hill Research Foundation.

Payne, C. (2008). *So much reform, so little change: The persistence of failure in urban schools.* Boston, MA: Harvard Education Press

perception of principal behaviors: A research study. Paper presented at the annual

Perrachione, B. A., Rosser, V. J., & Petersen, G. J. (2008 Fall) Why Do They Stay? Elementary Teachers' Perceptions of Job Satisfaction and Retention. *Professional Educator*, 32(2)

Petrilli, M. J. & Skull, J. (2011). *American Achievement in International Perspective*. Washington, D.C.: Fordham Institute Press

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Porter, A. C., Polikoff, M. S., Goldring, E., Murphy, J., Elliott, S. N., & May, H. (2010). Developing a psychometrically sound assessment of school leadership: The VAL-ED as a case study. *Educational Administration Quarterly*, 46 (2), 135-173.

Quaglia R., Marion S. F., McIntire W. G. (1991). The relationship of teacher satisfaction to perceptions of school organization, teacher empowerment, work conditions, and community status. *Education*, 112, 206-217.

Ravich, D. (2013) *Reign of Error: The Hoax of the Privatization Movement and the Danger to America's Public Schools*. New York: Vintage Books

Reimers, F.M., & Kanter, R.M. (2014). Education For The 21st Century. Harvard University Advanced Leadership Initiative Think Tank. 1-39. advancedleadership.harvard.edu/files/ali/files/2014_education_report_web.pdf?m=141209 1214

Richardson. M. Flanigan. J.. Lane. K.. & Keaster. R. (1992). Teacher perception of principal behaviors: A research study. Paper presented at the annual meeting of the Mid-South Educational Research Association in Knoxville. TN.

Rinehart , J. S., & Short, P. M. (1994). Job satisfaction and empowerment among teacher leaders, reading recovery teachers and regular classroom teachers. *Education* , 114(4), 570-580.

Rinehart, J. S., & Short, P. M. (1991). Viewing reading recovery as a restructuring phenomenon. *Journal of School Leadership*. 1(4), 379–399.

Robinson, V., Lloyd, C., & Rowe, K. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly*, *44*(5), 635-674.

Rogers, F. & Richarme, M. (2009). Decision Analyst. "The Honesty of Online Survey Respondents: Lessons Learned and Prescriptive Remedies". Retrieved from http://www.decisionanalyst.com/Downloads/honestyOnlineRespondents.pdf

Boglera, R., & Some, A. (2004). Influence of teacher empowerment on teachers' organizational commitment, professional commitment and organizational citizenship behavior in schools. *Teaching and Teacher Education,* 20, 277–289

Ross, J. A., & Gray, P. (2006). School leadership and student achievement: The mediating effects of teacher beliefs. *Canadian of Journal of Education, 29(3),*798-822.

Rotberg, I. C. (2006). Assessment around the World. *Educational Leadership*, 64(3), 58-63.

Ryan, J. (1998). Critical leadership for education in a postmodern world: Emancipation, resistance, and communal action. *International Journal of Leadership in Education, 1*(3), 257-278.

Schleicher, A. (2009). Seeing the United States Education System through the Prism of International Comparisons. *Middle School Journal,* 40(5), 11-17.

Schram, T. (2006) *Conceptualizing and Proposing Qualitative Research.* OH: Pearson Education

Schwartz, R. & Mehta, J. (2011). Finland: Superb Teachers – How to Get Them, How to Use Them*: Surpassing Shanghai: An Agenda for American Education Built on the World's Leading Systems* (M. Tucker, editor), Harvard Education Press

Sergiovanni, T. (2005). The Virtues of Leadership. *The Educational Forum*, 69(Winter), 112-123. Retrieved July 27, 2015, from http://www.scribd.com/doc/7375166/Sergiovanni-Thomas-Virtues-of-Leadership

Short, P. M. (1994a). Defining teacher empowerment. *Education*, 114(4), 488–492.

Short, P. M., & Rinehart, J. S. (1992). School Participant Empowerment Scale: Assessment of level of participant empowerment. *Educational and Psychological Measurement*, 54(4), 950-961.

Short, P. M., Greer, J. T., & Melvin, W. M. (1994). Creating empowered schools: Lessons in change. *Journal of Educational Research*, 32(4), 38–52.

Silins, H., & Mulford, B. (2004). Schools as Learning Organizations - Effects on Teacher Leadership and Student Outcomes. *School Effectiveness and School Improvement*, 3-4.

Simola, H. (2005) The Finnish miracle of PISA: Historical and sociological remarks on teaching and teacher education. *Comparative Education, 41(4), 455-470*

Sivanathan, N. & Fekken, C.G. (2002). Emotional intelligence, moral reasoning and transformational leadership. *Leadership & Organization Development Journal, 23(4), 198 – 204*

Skerrett, A. (2010). Of Literary Import: A Case of Cross-National Similarities in the Secondary English Curriculum in the United States and Canada. *Research in the Teaching of English*, 45(1), 36-58.

Sobe, N.W. & Kowalczyk, J.A. (2012). Exploding the Cube: Revisioning "Context" in the Field of Comparative Education. *Current Issues in Comparative Education, 16(1), 6-12.*

Sparks, S.D. (2014). Global Survey of Teachers Finds High Job Satisfaction but Cracks Under Surface. *Education Week.* Retrieved from http://www.edweek.org/ew/articles/2014/07/09/36satisfaction.h33.html

Spillane, J. P., & Sherer, J. Z. (2004). A Distributed perspective on school leadership: Leadership practice as stretched over people and place. Paper presented at the American Education Research Association Annual Meeting and Conference, San Diego, California

Starcher, S. L. (2006). The relationship between leadership practices of principals and student achievement (Doctoral dissertation). Retrieved April 9, 2016 from http://www.marshall.edu/etd/doctors/starcher-serena-2006-phd.pdf

Steiner-Khamsi, G. (2012). Understanding Policy Borrowing and Lending. Building Comparative Policy Studies. In G. Steiner-Khamsi & F. Waldow, eds., Policy Borrowing and Lending. *World Yearbook of Education 2012*. London and New York, pp. 3-17.

Sullivan, G. & Artino A (2013) Analyzing and Interpreting Data From Likert-Type Scales. *J Grad Med Educ* Dec; 5(4): 541–542.

Swackhamer, E.Y., Koellner, K., Basile, C., & Kimbrough, D. (2009). Increasing the Self-Efficacy of In-service Teachers through Content Knowledge. *Teacher Education Quarterly, 37(3), 63-78*

Sweetland, S.R., & Hoy, W.K.(2000). School characteristics and educational outcomes: Towards an organizational model of student achievement in middle schools. *Educational Administration Quarterly*. 36(5),703-729

Taylor, T. & Martin, B.N., Hutchinson, S., Jinks, M. (October-December, 2007) Examination of leadership practices of principals identified as servant leaders. *International Journal of Leadership in Education*. 10(4) 401-419.

The Associated Press-NORC Center for Public Affairs Research. (2013). *Parents' attitudes on the quality of education in the United States.* Chicago, IL: Tompson, T., Benz, J., & Agiesta, J.

The Gordon Commission. (2012) *Test-Based Accountability.* Princeton, NJ: Linn, R.

TIMSS and PIRLS International Study Center (2013) http://timssandpirls.bc.edu/

Townsend, T. Thinking and acting both locally and globally: new issues for teacher education. 2011. Journal of Education for Teaching: International research and pedagogy. Volume 37, Issue 2.

Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68, 202 - 248*

Tucker, Marc. (Ed.). (2011). *Surpassing Shanghai: An Agenda for American Education Built on the World's Leading Systems*. Cambridge, MA: Harvard Education Press.

Turgut, G. (2013). International Tests and the U.S. Educational Reforms: Can Success Be Replicated? *The Clearing House: A Journal of Educational Strategies*, 86(2), 64-73.

UNICEF Innocenti Research Centre (2012). Measuring Child Poverty: New league tables of child poverty in the world's rich countries. Innocenti Report Card 10, UNICEF Innocenti Research Centre, Florence

Vasquez Heilig, J., Jez, S. & Reddick, R. (2012). Is Texas leading its peers and the nation?: A Decadal Analysis of Educational Data. The Institute for Urban Policy Research and Analysis. University of Texas at Austin.

Veuglers, V. & O'Hair, J.M. (Eds) (2005), *School-University Networks and Educational Change*, Open University Press/McGraw-Hill, Maidenhead.

Visser, P. S., Krosnick, J. A., & Lavrakas, P. (2000).Survey research In H. T. Reis & C. M. Judd (Eds.), Handbook of research methods in social psychology. New York: Cambridge University Press.

Wall, R., & Rinehart, J. S. (1997). School-Based Decision Making and the Empowerment of Secondary School Teachers. Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28).

Wallace Foundation. (March 2009). Assessing The Effectiveness Of School Leaders: New Directions And New Processes. Retrieved from http://www.wallacefoundation.org/knowledge-center/school-leadership/principal-evaluation/Documents/Assessing-the-Effectiveness-of-School-Leaders.pdf

Whaley, K. W. (Oct 1994). Leadership and Teacher Job Satisfaction. *NASSP Bulletin*, 78(564), 46-50

Whitaker, D. S., & Moses, M. C. (1990). Teacher empowerment: A key to restructuring. Clearing House, 64 (2), 127-134.

Willms, J.D. (2004). Why the Difference? Variation in Reading Scores Among Canadian Provinces. *Education Canada, 45(1), 54-55.*

Woolfolk-Hoy, A. (1998). Complexity and coherence: An essay review of the Handbook of educational psychology. *Teachers College Record*, 100, 437-452.

York-Barr, J., & Duke, K. (2004). What do we know about teacher leadership? Findings from two decades of scholarship. *Review of Educational Research, 74*(3), 255-316.

Zepeda, S. J. (2007). Instructional Supervision Applying Tools and Concepts. Larchmont, New York, Eye on Education.

Zhao, Y. (2007). Education in the flat world: Implications of globalization on education. *Edge Magazine,* 2(4), 1-19.

Zhao, Y. (2014). *Who's Afraid of the Big Bad Dragon?: Why China Has the Best (and Worst) Education System in the World*. San Francisco: Jossey-Bass

Zimmerman, S. & Deckert-Pelton, M. (2003). Evaluating the evaluators: Teachers' perceptions of the principal's role in professional evaluation. National Association of Secondary School Principals. *NASSP Bulletin*, 87, 636; 28-37.