STARS

Electronic Theses and Dissertations, 2004-2019

2018

Interpersonal Behavior Traits and Their Relationship to Administrator-to-Teacher Feedback: A Quantitative Study

Robert Walker
University of Central Florida

University of Central Florida

STARS
Showcase of Text, Archives, Research & Scholarship

INTERPERSONAL BEHAVIOR TRAITS AND THEIR RELATIONSHIP
TO ADMINISTRATOR-TO-TEACHER FEEDBACK: A
QUANTITATIVE STUDY

by

ROBERT C. WALKER II
M.F.A. Virginia Tech, 2010
B.A. Rollins College, 2005

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Education
in the School of Teaching, Learning, and Leadership
in the College of Education and Human Performance
at the University of Central Florida
Orlando, Florida

Spring Term
2018

Major Professor: Barbara Murray

ABSTRACT

The purpose of this study was to explore what, if any, relationship exists between the interpersonal behavior traits held by administrators and the quality of the feedback they provide to teachers. The Interpersonal Behavior Survey (IBS) was used to develop interpersonal behavior profiles for all the school-based administrators from a moderately sized school district who consented to participate in the study. Additionally, the comments submitted as feedback to teachers by the participating administrators were reviewed and scored using a rubric.

Multiple regression analysis was performed to determine what, if any, relationship exists between the traits measured by the IBS and feedback quality. The IBS is divided into four scales: validity, assertiveness, aggressiveness, and relationship. These groupings were used to formulate the four research questions that guided this study: (1) what, if any, relationship exists between assertiveness traits and feedback quality, (2) what, if any, relationship exists between aggressiveness traits and feedback quality, (3) what, if any, relationship exists between relationship traits and feedback quality, and (4) what, if any, relationship exists between scores above the cut-off for reliability on any of the three validity scales and the quality of feedback given. No significant relationship was found to exist between any of the four IBS scale groupings and feedback quality; however, power analysis showed the lack of significance observed in this study could be due to the size of the population and not a true lack of significance. The study did find a significant relationship between age and years of experience in administration and feedback quality.

This study is valuable in that it contributes to the conversation regarding teacher effectiveness ratings, feedback, and sheds light on the role interpersonal behavior traits held by the administrator play in the feedback giving process. This study suggests there is reason to continue exploring the important role conflict avoidance may play in teacher evaluation and teacher effectiveness ratings.

This dissertation is dedicated to all the great educators who have guided me on my journey through this world. I thank you all for your wisdom, your generosity, your kindness, and your unwavering belief that I was capable of more than I realized. I would not be who and where I am today without each of you. Thank you.

ACKNOWLEDGMENTS

The completion of this doctoral journey would not have been possible without the encouragement and support of many, many people.

My sincere thanks to the two principals—Edward Jones & Andrew Jackson—who allowed me tremendous opportunities for career-growth and professional-learning while being understanding of the toll late nights of dissertation writing take. I hope, Mr. Jones, when my career comes to an end, I can walk away knowing I did it with as much class and integrity as you have modeled. And a special thank-you to Mr. Jackson, my mentor, you, sir, are a leader of rare passion and wisdom. I am and will continue to be thankful for the countless hours you have been willing to indulge my thinking aloud, to offer career advice, and, well, to be in all ways a true mentor and friend; thank you, Cadillac.

My thanks to my parents, Coleman and Sonya, who instilled in me, at an early age, the value of education. Thanks to my father for always modeling professionalism and dedication to a job, your work ethic is a fine inheritance. And more thanks than I can adequately express here, to my mother for modeling perseverance and unwillingness to ever give-in, even when the road got rough—everything I know about pushing through, I learned from you.

Many thanks to my dissertation committee. I have been truly fortunate to have the collective wisdom of Dr. Kenneth Murray, Dr. Lee Baldwin, Dr. Mary Alice Myers, and Dr. Barbara Murray guiding me along this journey. Additional thanks are due to Dr. Barbara Murray, my dissertation chair, without whom this study would not exist. You have truly inspired me, and I only hope the career ahead of me does justice to all you

taught me. I consider it my great good fortune to have the fine distinction of being the author of the last dissertation you chaired, I am humbled to carry that honor. Thank you, thank you, thank you.

And, many thanks to my colleagues in cohort six. I am so fortunate to have journeyed down this road with such a wonderfully thoughtful, kind, good-hearted, and diverse group of educators. I am thankful for all that I learned from each of you. I hope our life and career paths cross again.

Special thanks to Dr. Debra Pace, Superintendent of Osceola County Schools and Dr. Scott Fritz, Chief of Staff for Osceola County, for approving this study and helping to guide it during its infancy. I would not have reached this milestone without your support and am thankful for it.

And, finally, my deep thanks to my husband-to-be, Alexander Fritz, who rode the turbulent waters of deadlines and sleeplessness by my side, and who was always there to remind me "yes you can" when the self-doubt grew strong. I could not have done this without your love and support. I am a very lucky man, indeed. And, yes, it is now time for some much overdue traveling and us time for me and you.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1
INTRODUCTION

Background of the Study

Since the beginning of the modern accountability in education movement began—approximately 1983 with the publication of A Nation at Risk and reaching its legislative-pinnacle in 2001 when No Child Left Behind (NCLB) was enacted—there has been an emphasis on student proficiency and, more importantly, on the effectiveness of teachers to move their students toward reaching proficiency, especially in reading and mathematics.

Studies (Mela, 2013, Pace, 2015, Rafalski, 2015, and Butler, 2017) have shown that, when compared to the achievement of their students, the evaluation scores, or effectiveness ratings, of teachers are inaccurate, and they show signs of being inflated. As Mela (2013) states: "This means that [Brevard County] is still experiencing to a lesser degree; however, what Weisberg et al. (2009) called the Widget Effect where all teachers' classroom effectiveness was judged the same with no distinction being made between the good and the poor teachers" (p. 137). In order for a poor teacher's score to match that of a good teacher, there must be score inflation occurring. In short, school administrators who evaluate teacher effectiveness, principals and assistant principals, are inflating teacher effectiveness scores. This inflation becomes apparent when teacher effectiveness scores are compared to student achievement scores. The study that follows, by examining teacher effectiveness ratings and relevant administrator feedback, sought to arrive at an understanding of why this inflation occurs.

The School District of Osceola County is located to the south of Orlando, Florida in what is commonly referred to as "Central Florida." The district serves the cities of Celebration, Kissimmee, Poinciana, St. Cloud, Harmony, and Kenansville. The district has 52 schools and serves 61,736 students.

Since 2011, The School District of Osceola County, and indeed the entire state of Florida, has, as mandated by Florida statute 1012.34, rated their teachers as highly effective, effective, needs improvement, developing (a rating only available for teachers in their first three years of service), or unsatisfactory. Florida statute 1012.33 provides that any teacher receiving two consecutive ratings of unsatisfactory, two ratings of unsatisfactory in a three-year period, or three consecutive ratings of needs improvement or a three-year combination of needs improvement and unsatisfactory can be terminated for failing to perform the duties of their job adequately. This statute prevents students in the state of Florida from being subjected to poorly performing teachers year after year by giving educational leaders in the state a legal means to unburden their schools of teachers who are not providing effective instruction to their students. However, given the inflation of teacher effectiveness ratings, the mechanism of this statute allowing school leaders to dismiss ineffective teachers is rendered inconsequential, because few ineffective teachers are actually receiving ratings that reflect their ineffectiveness. As such, the stipulations of the statute are never triggered and the ineffective teacher—thanks to their administrator giving them inflated ratings—continues to provide ineffective instruction to students.

For the 2012-13, 2013-14, and 2014-15 school years the percentage of teachers rated as either effective or highly effective never failed to exceed 97% statewide and 95%

in Osceola County. Statewide, teacher evaluations were highest for 2014-15, 98.4% effective or highly effective, and Osceola County's administrators rated their teachers highest in 2012-13, 99% effective or highly effective. Logic creates the expectation that a cohort of teachers who have, for three consecutive years, been rated effective or highly effective at a rate of 97% or better, statewide, or 95% or better, in Osceola County, would produce students whose achievement scores were indicative of having received effective or highly effective instruction; however, this is not the case. For both the 2014-15 and 2015-16 school years, The School District of Osceola County saw 50% or less of its students reach proficiency in either math or reading. The performance statewide was not much better with reading proficiency stagnating at 52% for both years, and math proficiency increasing only one percentage point from 56% in 2015 to 57% in 2016. These numbers clearly illustrate a chasm between the ratings of teacher effectiveness, as determined by administrators, and the reality of their effectiveness as illustrated by student achievement. See table 1 below for a full breakdown of state and district proficiency across multiple subject areas. See table 2 below for a breakdown of teacher effectiveness ratings in the district and across the state.

Table 1. State & District percent of students proficient by year and subject.

Subject	State 2015	Osceola 2015	State 2016	Osceola 2016
ELA 3-10	52	49	52	48
Math 3-8	56	50	57	48
Algebra 1 4-12	56	67	54	50
Geometry 5-12	53	41	51	51
Algebra 2 5-12	36	29	40	31
Biology 1 6-12	65	63	64	69
Civics 3-12	65	61	67	70

Table 2. State & district teacher effectiveness ratings by year.

Rating	State 2014	Osceola 2014	State 2015	Osceola 2015
Highly Effective	41.9	64.9	37.5	45.7
Effective	55.7	33.6	60.9	49.5
Needs Improvement	1.4	0.9	1	1.8
Developing	0.7	0.6	0.5	0.9
Unsatisfactory	0.3	0	0.2	2.2

What is of particular interest in this study is the the low rate of student performance despite the high ratings of teacher effectiveness. Over the course of the three school years leading to the student performance delineated above—2012-13, 2013-14, and 2014-15— teacher effectiveness averaged 97.9% of all teachers in the state being either effective or highly effective (60.7% effective and 37.2% highly effective). The teacher ratings in The School District of Osceola County were similar: 97.6% highly effective or effective (50.3% effective and 47.2% highly effective). During this same span of time only 1.5% of teachers, statewide, were rated as needing improvement or unsatisfactory. Only 1.8% of teachers in Osceola County were rated as needing improvement or unsatisfactory during the same period of time. The crucial question here is: How is it possible to have sustained better than 95% of teachers being effective or highly effective, yet have student proficiency on state assessments in the neighborhood of

4

50%? These numbers are incompatible; they defy logic. One cannot be both "effective" yet fail to meet or arrive near one's set goal.

While much has been made of the various feedback and evaluation systems— Marzano and Danielson to name two—that are available to school districts, this study posits that simply looking at the particular feedback protocol as the explanation for the disparity between teacher rating and student performance is a gross oversimplification that risks missing the true root cause of the discrepancy. What is at issue is the administrators making use of the feedback tool, not the tool itself.

Thus far the discussion has touched exclusively on teachers and students; however, there is another crucial party in all of this: The administrator. The administrator, or instructional leader, is crucial to this conversation because he or she is the person completing the evaluations and rating the teachers. As John Brooks Slaughter (2012) states, "…a good leader is one who demonstrates the capacity to bring about positive change…" (p. 85). Manz and Sims (1991) further this definition when they state, "In many modern situations, the most appropriate leader is one who can lead others to lead themselves" (p. 213). How do instructional leaders lead teachers to lead themselves and bring about positive change? The answer here is akin to how teachers accomplish similar goals with students: Feedback. As Hattie (2009) notes, "…feedback [is] among the most powerful influences on achievement" (p. 173). Feedback, of course, does not originate in a vacuum, nor does the use of any feedback protocol remove the human element from the feedback ecosystem. The giving of feedback is essentially an interpersonal communication, and, as such, will be impacted by the interpersonal behavioral traits of

the individual offering the feedback. Because the giving of feedback requires interpersonal interaction, if the root flaw in the feedback is caused by an excess or deficiency in certain interpersonal behavioral traits, no feedback protocol can be successful until those excesses or deficiencies are addressed.

As the numbers presented on tables 1 and 2 above show, teacher ratings and student achievement are horribly out of alignment: While teachers continue to be rated as effective or highly effective at a very high rate, the students they instruct continue to not demonstrate the academic gains that would evidence such effective instruction. The reality of teacher effectiveness rating inflation is an easily observed phenomenon, what is less clear is why the phenomenon continues to occur. By examining the interpersonal behavioral traits of those tasked with providing feedback to and evaluating teachers, in The School District of Osceola County, this study aimed to arrive at an understanding of why this phenomenon occurs.

Statement of the Problem

While the fact that teacher effectiveness ratings, when compared to student achievement, are inflated is well established; to date, limited research has been conducted to determine what is causing this inflation to occur. This inflation renders the evaluation of teachers ineffective. To date, no study has sought to examine the effect interpersonal behavior traits, such as aggressiveness, assertiveness, conflict avoidance, and other relationship traits have on the feedback given to teachers by the administrators who assess them. Rather, researchers have focused, nearly exclusively, on constructing rubrics

6

for scoring classroom instruction, and, in a few instances, the impact of interpersonal relationships—and the behaviors that govern them—between teachers and their students. Given the crucial role administrators play as instructional leaders tasked with cultivating quality instruction across the campuses they serve, this study is vital to determining the role interpersonal behavioral traits play in the administrator-to-teacher feedback ecosystem, because, if the root cause of inflated and thus ineffective feedback and ratings lies in interpersonal behavior flaws or shortcomings, then no feedback protocol can be successful until these underlying interpersonal behavioral concerns are addressed. This study addresses the impact of the person using the tool, rather than simply assessing the tool, as previous studies have done.

Purpose of the Study

The purpose of this study is to, using the Interpersonal Behavior Survey (IBS), develop interpersonal behavior profiles for all the teacher-assessing administrators in a district, and then to analyze the feedback provided by those administrators to determine if any relationship between specific interpersonal behavioral traits, such as conflict avoidance, and the quality of administrator feedback exists.

Significance of the Study

This study is significant because it sought to explain something that heretofore has remained unexplained: Why are the vast majority teachers in Florida, and in The School District of Osceola County specifically, continually rated as effective or highly effective, while the students of the state and district alike continue to perform in ways

7

that do not demonstrate having received effective instruction? Unlike any prior studies, this study did not seek to find fault in any particular feedback protocol, nor did it seek to advocate for one protocol over all others; however, it did seek to examine and understand the importance of interpersonal behavioral traits in the feedback process. No existing study has sought to explain the discrepancy in teacher rating and student performance by examining the interpersonal behavioral traits of the administrators providing the feedback.

This study sought to expand the understanding of the process of communicating feedback, and, more importantly, highlighted specific interpersonal behavior traits that when in excess or deficiency are problematic for the administrator attempting to improve teacher performance with his or her feedback. The study sought to, for example, show that an excessive level of conflict avoidance corresponds with giving less specific actionable feedback. Feedback of this type is, of course, of less use to the teacher and thus less likely to lead to any improvement in instruction and, in turn, any improvement in student performance.

The study employed the Interpersonal Behavior Survey (IBS) (Mauger and Adkinson, 1980) to develop interpersonal behavior profiles for each of the administrators in the study. The IBS was developed to detect excesses and deficits in aggressiveness and passivity, along with specific relationship traits (Mauger & Adkinson, 1980, p. 1). The IBS is composed of seven scales under aggressiveness, eight under assertiveness, and three under relationship. The seven aggressiveness scales are: general aggressiveness, hostile stance, expression of anger, disregard for rights, verbal aggressiveness, physical

aggressiveness, and passive aggressiveness. The eight assertiveness scales are: general assertiveness, self-confidence, initiating assertiveness, defending assertiveness, frankness, praise, requesting help, and refusing demands. The three relationship scales are: conflict avoidance, dependency, and shyness. The IBS also employs three scales to measure the validity of the responses provided; the three validity scales are: denial, infrequency, and impression management. A thorough definition of each scale is provided in the definition of terms section of this study.

If the study had identified certain interpersonal behavioral traits that are problematic and others that are beneficial in the feedback giving process, then it would have enhanced the ability of school district level leadership to provide training to hone the beneficial traits and diminish the impact of the problematic ones. This would allow the root cause of the discrepancy to be addressed rather than simply changing from one feedback or evaluation protocol to another; because, as the numbers show, this discrepancy in teacher rating to student performance exists across the state despite the fact that districts are using a variety of different feedback and evaluation protocols.

Hattie (2009) has already shown that the relationship between teacher and student has a measurable impact on student achievement, by examining the interpersonal behavioral traits of administrators and the performance of the teachers they supervise, this study sought to demonstrate a similar power of relationship between administrator and teacher.

<center>Definition of Terms</center>

School District of Osceola County (SDOC): A moderately sized school district located in Osceola County, Florida.

School-Based Administrator: Principals and assistant principals

Interpersonal Behavior Survey (IBS): First published in 1980 and developed by Paul A. Mauger, Ph.D., David R. Adkinson, Ph.D., Suzanne K. Zoss, Ph.D., Gregory Firestone, Ph.D., and J. David Hook, MA, the IBS measures various dimensions of assertive and aggressive behavior on the following scales: general aggressiveness, conflict avoidance, frankness, hostile stance, dependency, praise (giving/receiving), expression of anger, shyness, requesting help, disregard for rights, general assertiveness, refusing demands, verbal aggressiveness, self-confidence, denial, and physical aggressiveness.

Validity Scales: The group of scales (denial, infrequency, and impression management) on the IBS that reflect test-taking attitudes. The respondent's attitude toward testing in general and the IBS in particular affects scores on the assertive, aggressive, and relationship scale. A score above 70T on any of the validity scales, renders the profile uninterpretable.

The Denial Scale (DE): This scale indicates a hesitancy to admit to common but socially undesirable weaknesses and feelings. This includes such behaviors as making fun of others, swearing, and procrastinating.

The Infrequency Scale (IF): This scale indicates the tendency to endorse items that less than ten-percent of the normative sample endorsed. Such items were selected

<center>10</center>

solely on the basis of frequency of endorsement, so that the scale spans a variety of content areas.

The Impression Management Scale (IM): This scale measures the degree to which impression management plays a part in a person's responses to IBS items. This scale was derived by asking college students to take the IBS twice. On the second administration they were instructed to respond in such a manner as to create a favorable impression of themselves. Items that changed significantly comprise the IM scale.

The General Aggressiveness, Rational Scale (GGR): This scale measures the general response class of aggressiveness over a wide variety of item content including aggressive behaviors, feelings, and attitudes.

The Hostile Stance Scale (HS): This scale measures an antagonistic orientation toward other people, a view of the world that justifies aggression in order to get ahead in life or to protect oneself.

The Expression of Anger Scale (EA): This scale is an indication of the tendency to lose one's temper and express one's anger in a direct, forceful manner.

The Disregard for Rights Scale (DR): This scale measures the tendency to ignore the rights of others in order to protect oneself or to gain an advantage.

The Verbal Aggressiveness Scale (VE): This scale gives an indication of the using of words as weapons by doing such things as making fun of others, criticizing, and putting others down.

The Physical Aggressiveness Scale (PH): This scale reflects the tendency to use or fantasize using physical force.

The Passive Aggressiveness Scale (PA): This scale measures behaviors that indicate indirect or passive expressions of aggressiveness. Such behaviors as stubbornness, negativism, procrastination, and complaining are sampled by this scale.

The General Assertiveness, Rational Scale (SGR): This scale is a general measure of assertiveness. The item content covers a broad range of assertive behaviors.

The Self-Confidence Scale (SC): This scale measures the expression of positive attitudes about one's self and the expression of self-assurance. It should be emphasized that this is self-confidence in one's interpersonal relationship skills. A person can have a negative self-image or not like himself or herself very much, yet still score high on this scale. It is also possible for a person with a positive self-image to score low if he or she feels inadequate in social relationships, especially those of a personal nature.

The Initiating Assertiveness Scale (IA): This scale measure is an indication of leadership potential and the tendency to take an ascendant role in groups. Behaviors sampled by this scale include a high frequency of verbal participation in public settings and a propensity to make suggestions and express one's opinions.

The Defending Assertiveness Scale (DA): This scale reflects behaviors related to standing up for one's rights.

The Frankness Scale (FR): This scale samples the willingness to clearly communicate one's true feelings and opinions even though these expressions may be unpopular or may cause a confrontation with others.

The Praise Scale (PR): This scale reflects one's degree of comfort in giving and receiving praise.

The Requesting Help Scale (RE): This scale measures the willingness to ask for reasonable favors and help when they are legitimately needed.

The Refusing Demands Scale (RF): This scale indicates the willingness to say "no" to unreasonable or inconvenient demands from others.

The Conflict Avoidance Scale (CA): This scale is modeled on Bach's (Bach & Wyden, 1968) notion of fight phobic behaviors. Persons scoring high on this scale tend to evade open disagreement or conflict with others.

The Dependency Scale (DP): This scale indicates the degree to which a person is dependent on others. Behaviors tapped by this scale include relying on others for help in decision making, feelings of powerlessness and helplessness, fear of losing the support of others, and attention seeking.

The Shyness Scale (SH): This scale samples social behaviors such as friendliness, participation in social events, and the enjoyment of social interaction. People who score high on this scale are socially introverted and feel more comfortable socializing with family members or a small circle of close friends than with larger groups.
All IBS scale definitions adapted from Mauger & Adkinson (1980).

Theoretical Framework

The German philosopher, economist, scientist, and sociologist Karl Marx offers the best theoretical frame through which to understand the true lineage and theoretical underpinnings of this study with his foundational thinking in the area of conflict theory. Marx posits that the bourgeois—the owners of capital, the ruling class—are in a constant

13

state of conflict with the proletariat—the working class (Mannion, 2002, p. 143). Marx "believed that economic relationships were of primary importance, and the conflict between classes was an inevitability, due to the chasm between the haves and have-nots" (Mannion, 2002, p. 143). Furthermore, Marx argues that "social order is maintained by domination and power, rather than consensus and conformity. According to conflict theory, those with wealth and power try to hold on to it by any means possible, chiefly by suppressing the poor and powerless" ("Conflict theory", 2018). This is the crux of Marx's theory: Those with the resources and power seek to maintain their status by suppressing those without—leading to conflict between the two. Furthermore, Marx asserted that "the majority of people toiled with little reward while the upper classes reaped the fruits of their labor" (Mannion, 2002, p. 144). According to a concise breakdown of Marx's conflict theory, it has:

> …been used to explain a wide range of social phenomena, including wars and revolutions, wealth and poverty, discrimination and domestic violence. It ascribes most of the fundamental developments in human history, such as democracy and civil rights, to capitalistic attempts to control the masses rather than to a desire for social order. The theory revolves around concepts of social inequality in the division of resources and focuses on the conflicts that exist between classes. ("Conflict Theory", 2018)

Additionally, Marx's theory has been used as a means to understand more recent events such as the the 2008 financial crisis. Cairns & Sears (2010) note that the same governments that claimed to lack sufficient funds for things like universal healthcare

programs were able to offer billions of dollars to the banking industry in the form of bailouts. This, they explain, is evidence of Marx's basic premise in action.

What does Marx's theory have to do with teachers and their assessing administrators? Everything. Schools, in many ways, are micro-societies. They exist as complete societies within other societies. In the society of school there exists the social stratification that exists in all societies: the bourgeoisie, who control the resources and power, and the proletariat, who lack both resources and power. In the context of a school, the administrators—who are often better compensated for their work and control decision making authority—are the bourgeoisie, the teachers are the proletariat. While invoking Marx's theory in what may seem to be the ultimate edifice of democratic ideals, the school house, may be a cause for discomfort for some, it is appropriate.

To further extend Marx's theory into the school setting, think of administrators as controlling access to the ultimate resource: Jobs. According to Florida state statutes, there are multiple combinations of repeated poor evaluations that will cause a teacher to lose their job. Furthermore, the administrators control access to desirable teaching assignments: Who teaches Advanced Placement (AP), Dual Enrollment (DE), or honors courses are all decisions made and controlled by administrators. This analogous understanding of Marx's theory in schools is not only applicable to high schools, where course offerings are more varied, but also to lower grades—the decision regarding the placement of a notoriously ill-mannered, ill-tempered, ill-prepared third grader is also one made by the school's bourgeoisie. No, these things do not constitute the incitement of full scale war, which is often the case with Marx's conflict theory, they do, however, lead

to conflict. Evidence of this conflict can be overheard in any teacher's lounge, school parking lot after dismissal, local watering hole where teachers congregate, or, in the modern context, on the pages of social networking sites such as Facebook or Twitter.

Tensions, as they are often want to do, will in some cases rise until there is palpable near-violent, if-not-violent, relations between teachers and their administrators. This inherent tendency toward conflict is precisely what Marx sees as an inevitability. The focus of this study, interpersonal behavior traits and feedback, tie directly to Marx's theory at work in schools because behavioral traits are evidence of this conflict manifest in how administrators engage with their teachers. Because of the inherent power and resource imbalance, Marx would argue that the teacher-administrator dynamic is one which will be inherently fraught with conflict. By placating the teacher with an inflated effectiveness rating, the administrator is able to quell any potential uprising and pacify the proletariat. Giving these inflated ratings costs the administrators nothing in terms of resources and power. These inflated ratings represent the path of least resistance and sure-fire way to consolidate support and maintain power; few are likely to rebel against an authority that constantly tosses down affirmations. Inflated ratings are a means by which the bourgeoisie—the administrators—pacify the proletariat—the teachers—and thus continue to maintain control of power and resources; continually coming out ahead in the ongoing conflict Marx ascribes as an inevitability of the human condition.

Research Questions

Given that the purpose of this study is to determine what, if any, relationship exists between interpersonal behavior traits and the quality of administrator-to-teacher feedback, the research questions that guide this study were developed around the interpersonal behavior traits measured by the IBS. The IBS groups these behaviors into four groups: assertiveness, aggressiveness, relationship, and validity. These groupings were used as the foundation upon which the research questions were constructed.

1. What, if any, relationship exists between the assertiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

2. What, if any, relationship exists between the aggressiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

3. What, if any, relationship exists between the relationship scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

4. What is the relationship between scores above 70 on the Denial, Infrequency, and Impression Management scales—all scores that render the aggressiveness, assertiveness, and relationship portions of the respondent's survey unreliable— and the feedback provided by the administrators with those scores?

<center>Delimitations</center>

This study will be delimited by the following factors:

1. Only school-based administrators, principals and assistant principals, who directly supervise teachers were assessed. No district level administrators were asked to complete the IBS.

2. Only administrators in Osceola County Public Schools were assessed.

3. Only administrators who have held the position of principal or assistant principal prior to the 2016-17 school year were assessed.

4. Only administrators in attendance for the IBS administration sessions (May 9, 2017 for principals and a preparing new principals session on May 18, 2017 for assistant principals) were assessed.

5. Only administrators who consented to participate in the study were assessed.

6. Administrators and teachers from Osceola High School were excluded from the study as requested by The School District of Osceola County.

<center>Limitations</center>

This study will be limited by the extent to which the following parameters will be applied:

1. The study is limited by the ability of study participants to correctly mark their IBS survey responses on the answer sheet.

2. The study is limited to examining written feedback provided by administrator's in the district's online teacher evaluation platform. No face-to-face, telephone,

email, or other forms of feedback can be examined within the parameters of this

study.

3. This study is limited by the low rate of participation by administrators in the study

population.

4. This study is limited to administrators in Osceola County. Generalization to other

populations is limited.

Overview of Methodology

Research Design

A quantitative study will be conducted, using the IBS, to determine the

relationship between the interpersonal behavioral traits an administrator possess, and the

level to which he or she possesses them, effects the quality of the feedback he or she

provides to the teachers he or she supervises. The study sought to unearth the root cause

for why, despite the introduction of a variety of different feedback protocols across the

state, a sizable discrepancy continues to exist between student achievement and teacher

effectiveness ratings, highlighting the fact that administrators across the state, and in

Osceola County specifically, are inflating the effectiveness rating of the teachers they

evaluate.

The IBS was used to measure the interpersonal behavioral traits listed below for

each school-based administrator. This allowed the researcher to develop an interpersonal

behavior profile for each school-based administrator in the district.

Interpersonal Behavior Traits to be Measured:

- Denial

- Infrequency

- Impression Management

- General Aggression

- Hostile Stance

- Expression of Anger

- Disregard for Rights

- Verbal Aggressiveness

- Physical Aggressiveness

- Passive Aggressiveness

- General Assertiveness

- Self-Confidence

- Initiating Assertiveness

- Defending Assertiveness

- Frankness

- Praise (Giving/Receiving)

- Requesting Help

- Refusing Demands

- Conflict Avoidance

- Dependency

- Shyness

A rubric (Rafalski, 2015) was used to assess and categorize the feedback given by administrators. The rubric rates feedback on a scale of 1 to 7. One on the scale is no feedback given. Two on the scale is unrelated feedback. Three on the scale is a recount of classroom events; this may include information regarding what is lacking in the lesson but does not include any specifics for improving instruction. Four on the scale is a general affirmative statement. Five on the scale is a reflective question. Six on the scale is for standardized feedback, or boilerplate feedback that can be pasted from the teacher observation platform. Seven on the scale is for specific targeted feedback.

IBM's SPSS statistical analysis software was used to perform the quantitative analysis in the form of multiple linear regression to ascertain the relationship between IBS scores and feedback provided.

Population

All 131 of the principals and assistant principals employed by The School District of Osceola County were invited to participate in this study. 54, or 41%, of the 131 school-based administrators in the district opted to participate in this study. Tables 3, 4, and 5 below provide demographic breakdowns of study participants. It should be noted here that 3 participants declined to provide their years of experience in administration when responding to the IBS survey, and 3 participants also declined to provide their gender on the IBS response sheet. Additionally, 5 study participants opted to not enter their age on the IBS answer sheet.

Table 3. Study participants by gender.

Male	Female
18	33

Table 4. Study participants' years of experience in administration.

3 or less	4 to 6	7 to 9	10 to 12	13 or more
5	11	10	4	20

Table 5. Study participants' age.

Under 40	41-45	46-50	51-55	56-60	61-66
3	19	10	8	5	4

Data Collection

The researcher used the Interpersonal Behavior Survey (IBS) to develop interpersonal behavior profiles for all 54 administrators in The School District of Osceola County who opted to participate in this study. The researcher used multiple linear regression analysis to determine if any statistically relevant relationship between various personality traits and the quality of feedback given by a person possessing those traits exists.

Data Analysis

Quantitative data from the IBS and from the feedback rubric scores were analyzed using both Pearson correlation and multiple linear regression analysis to determine the relationship between interpersonal behavior traits possessed by an administrator, and the level to which he or she possesses them, and the feedback provided by that administrator to the teachers he or she supervises. Specifically, IBM's SPSS was used to conduct both

22

types of analysis to determine the relationship between the IBS scores and feedback scores.

Organization of the Study

This research study is presented in five chapters. Chapter 1 includes the background of the study, the statement of the problem, purpose of the study, definition of terms, theoretical framework, research questions, limitations, delimitations, overview of methodology, and the organization of the study. Chapter 2 presents a review of the literature relevant to the study. Chapter 3 describes the methodology used for this research study including an introduction, the research design, the selection of study participants, instrumentation, data collection, data analysis, and a summary. Chapter 4 presents the study's findings. Chapter 5 provides a summary of the entire study, discussion of the findings, implications of the findings, recommendations for further research, and a conclusion.

CHAPTER 2
REVIEW OF LITERATURE

Introduction

At the time of the present study, the use of the Marzano teacher evaluation system and the accompanying teacher effectiveness ratings are relatively new additions to the education landscape. These additions are founded in the belief that a significant direct relationship exists between student academic achievement and teacher effectiveness. In Florida, teacher effectiveness ratings are required by law. Florida statute 1012.34 stipulates that teacher effectiveness ratings "differentiate among four levels of performance as follows: 1. Highly effective. 2. Effective. 3. Needs improvement or, for instructional personnel in the first 3 years of employment who need improvement, developing. 4. Unsatisfactory." Teacher ratings are so important in the state of Florida that state statute 1012.33 lists "three consecutive annual performance evaluation ratings of needs improvement or combination of needs improvement and unsatisfactory" among the few just causes for termination of a teacher in the state. Florida is not unique in its requirement for teacher effectiveness ratings, nor is it unique it its desire for effective teachers. In 2011, as a requirement for receiving federal Race to the Top dollars, former United States Secretary of Education Arne Duncan stipulated that school districts begin reporting the percentage of their teachers in each effectiveness rating category (Donaldson, 2009). Secretary Duncan's call for an accounting of teacher effectiveness ratings was a direct decedent of the ideas put forth in the 1983 study *A Nation at Risk.* The study, conducted and published by the Reagan Administration's Education

Department, put forth the idea of reforming—strengthening— teacher evaluation as means for combating "the rising tide of mediocrity [in American education]" (Donaldson, 2009, p. 4).

The idea of evaluating the effectiveness or performance of workers is neither new nor unique to the education context. Dominant belief on the topic is that the use of the term performance appraisal came into use sometime around the end of World War II (Wiese & Buckley, 1998). Among the primary goals and benefits of any performance appraisal or effectiveness rating system is to identify gaps, weaknesses, and deficiencies in an employee that cause inefficiency and poor worker performance; once identified supervisors would then work to decrease these gaps, weaknesses, and deficiencies (Manasa & Reddy, 2009). The primary means for a supervisor to aid in decreasing weaknesses and deficiencies, and close gaps is by providing feedback to the employee, be it a teacher or any other worker. Hattie (2009) states that feedback is among the most powerful tools for achieving improvement; thus it can be understood that administrator-to-teacher feedback is a key component in improving teacher performance.

Despite the legislative mandates requiring effectiveness ratings and the teacher-effectiveness improving goal of teacher evaluation over the past decade, several studies (Mela, 2013, Pace, 2015, Rafalski, 2015, and Butler, 2017) have shown that when compared to student performance, teacher evaluations are inaccurate due to inflation. Additionally, little evidence exists to support the idea that administrator evaluations reliably identify effective and ineffective teachers when student performance is considered (Weisberg, et al,. 2009).

In the pages that follow this chapter presents the context for conducting further research into why, despite the intentions of teacher effectiveness ratings and the power of feedback, the evaluations of teachers by administrators continue to show evidence of inflation and ineffectiveness. As Lunenburg and Irby (2008) state, this chapter "will provide the basic rationale for [the] research" (p. 137). As the research questions for this study focus on the potential role interpersonal behaviors, including conflict avoidance, play in the inflating of teacher evaluation scores, this chapter will explore studies relevant to teacher evaluation as well as those relevant to communication, interposal behaviors, conflict avoidance, and moral leadership.

The researcher made use of the UCF Library One Search online tool to find resources to support the discussion related to the research questions of this study. The researcher also made use of the ERIC library, previously published dissertations and theses, Google Scholar, numerous and varied internet sources, scholarly journals, periodicals, books, published reports, and statutes from the state of Florida. Upon completion of the gathering of relevant documents and other materials, the researcher divided the information into six topics for discussion: (a) A discussion of the observed Lake Wobegon Effect (b) Feedback, (c) The Marzano casual teacher evaluation system, (d) Communication, (e) A brief overview of interpersonal communication and interpersonal behaviors as a subset there of, (f) A discussion of relevant findings relating to the role and influence of conflict avoidance, and (g) a discussion of moral leadership.

<u>The Lake Wobegon Effect</u>

In 1985 Garrison Keillor published the book *Lake Wobegon Days*. The characters from the fictitious town in which the novel is set later become regular components of Keillor's *A Praire Home Companion,* a regularly occurring NPR program. What makes the characters from Lake Wobegon worth discussion is how uniquely exceptional the town's inhabitants are. Writing in the New York Times, Keillor (2001) states, "Twenty-five years ago, for amusement, I invented a small town where the women are strong and the men good-looking and the children all above average…." What began as a fanciful, fictitious lark for Keillor has evolved from pop-cultural touch-point to a means used by social scientists to explain a unique human phenomenon: The tendency to perceive ourselves as always above average.

Social psychologist David G. Myers uses the Lake Wobegon effect as a means of labelling and understanding human behavior. Myers (2000) provides several examples of the Lake Wobegon effect in action: "Most businesspeople see themselves as more ethical than the average businessperson. Ninety present of business managers rate their performance as superior to their average peer. Most drivers—even most drivers who have been hospitalized for accidents—believe themselves to be safer and more skilled than the average driver" (p. 168). One study of 829,000 high school seniors preparing to enter college found that 60 percent rated themselves in the top 10 percent in their ability to get along with others, and 25 percent of the respondents rated themselves as being in the top 1 percent (Myers, 2000). One study (Hayes and Schaefer, 2009) used the Lake Wobegon effect as a means of explaining rising CEO salaries: No firm wants to admit to having a

less-than-average or even an average CEO so no CEO pay package can be below or even at average, every new package continually ups the average in order to bestow fair market compensation upon all the above average CEOs. Another study (Wolf and Wolf, 2013) uses the Lake Wobegon effect as a means of explaining overly optimistic framing for treatment plans developed for elderly cancer patients. At The University of Minnesota the Lake Wobegon effect has become part of the discourse over grade inflation and grade disparities from one college to another—62 percent of the university's education majors received grades above C, while the same was true of only 28 percent of its engineering students; one Engineering professor decries the grade inflating caused by the Lake Wobegon effect: "We've abrogated our responsibility as educators to distinguish between excellent, average, and good students" (Berrett, 2012). Substitute "teachers" for "students" and this professor's statement captures what is occurring in teacher evaluation today. The Lake Wobegon effect is, clearly, both a real and wide-ranging phenomenon of the human condition.

Kruger (1999) provides one definition for the Lake Wobegon effect, writing that "like the inhabitants of Garrison Keillor's fictional community of lake Wobegon, most people appear to believe that their skills and abilities are above average….This tendency engenders the oft-documented *above-average effect…"* (p. 221). Wheeler and Haertel (1993) take this understanding of the Lake Wobegon Effect as inflation of self-appraisal and add to it to include inflation of appraisal of others and groups: "a phenomenon in which most individuals or groups perform above average" (p. 82).

The Lake Wobegon effect is observable in education not only in the context of inflated university student grades, but also in the skewed appraisals of the abilities of our nation's educators. Studies (Ehrgott, Henderson-Sparks, and Sparks, 1995 and Fuhr, 1993) have estimated teacher incompetence to be around five percent across the profession; however, these studies note that teacher dismissal rates are far less than one percent across the profession. Tucker (1997) asks if there is "a discrepancy [between incompetence rates and dismissal rates] because we have realized the idyllic existence of Garrison Keillor's Lake Wobegon or because we are ignoring the serious problem of teacher incompetence?" (p. 104). The gap between the percentage of incompetent teachers and teachers being dismissed along with the high percentage of teachers being rated effective or highly effective despite student achievement measure numbers that fail to reflect such high levels of teacher competence are both evidence of the Lake Wobegon Effect in action.

This inflation of teacher ratings, The Lake Wobegon Effect, is not constrained to any particular school, district, or state. It has been documented across the United States. Student achievement outcomes do not substantiate these ratings. In Pennsylvania, 98.2% of teachers were given a perfect score on their evaluation (Langlois & Colarusso, 1988). In Florida Mela (2013) found that 99% of 1,138 Brevard County teachers were rated effective or highly effective. Two years later, again in Florida's Brevard County, Pace (2015) found that 92.5% of all the teachers were rated as effective or highly effective. Again, student performance outcomes from the state and district do not substantiate the high effectiveness ratings for the teachers. Again in Florida, Butler (2017) provides an

29

example of the Lake Wobegon Effect in the extreme: of the 528 Florida teachers comprising the lowest 10% of teachers as rated by Florida's Value Added Model (VAM), 43.2% were rated highly effective and 51.1% were rated effective. 94.3% of the lowest performing teachers in Florida were rated as either effective or highly effective by their administrator. When the Lake Wobegon effect is in effect even the worst teachers are nearly all above average.

As Donaldson (2010) notes, "time and again, analyses of summative evaluation ratings of teachers show that the vast majority of teachers in any school district, or state are rated above—sometimes well above—average" (p. 54). This reoccurring phenomenon has not gone unnoticed or unrealized by those guilty of the inflation: One study of Chicago administrators noted that 56% of principals admitted to assigning evaluation scores higher than what was warranted (Donaldson, 2009). In the same study, Donaldson (2009) noted that only 46% of principals nationwide gave their school a rating of excellent. Donaldson (2009) concludes that "any school…is more likely to employ more under-performing teachers than its evaluations ratings suggest. In fact, principals and teachers believe that teachers are less effective than evaluations ratings would indicate" (p. 2).

<center>Feedback</center>

Feedback is a central component of this study. While the study focuses on an aspect of administrator-to-teacher feedback, in reviewing the relevant literature the researcher considered feedback from any supervisor to subordinate. Given the rather extensive coverage of feedback in the available literature, this section of the literature will approach feedback in three parts: attempting to arrive at a definition of feedback from the literature, exploring the use of and need for feedback as covered in the literature, and, finally, an overview of some worthwhile qualifiers regarding feedback observed in the relevant literature.

Perhaps the best place to start in the effort to achieve a satisfactory definition of feedback is with the definition provided by Hattie & Timperly (2007):

> In this review, feedback is conceptualized as information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding. A teacher or parent can provide corrective information, a peer can provide an alternative strategy, a book can provide information to clarify ideas, a parent can provide encouragement, and a learner can look up the answer to evaluate the correctness of a response. Feedback thus is a "consequence" of performance. (p. 81)

For the purposes of this study, the teacher in both definitions would be the administrator and the student would be the classroom teacher. Sadler (1989) provides a similar definition of feedback, writing that "feedback is a key element in formative assessment, and is usually defined in terms of information about how successfully something has

<center>31</center>

been or is being done" (p. 120) Shute (2008) harkens back to an old Russian proverb as a means of understanding feedback: "It is not the horse that draws the cart, but the oats" (p. 153). Here the feedback servers as a motivating force. Shute (2008) goes on to explain that the primary goal of feedback is to "increase student knowledge, skills, and understanding in some content area or general skill" (p. 156). Shute (2008) goes on to assert that feedback should provide two types of information to the learner: verification and elaboration. In this context verification is understood as "the simple judgment of whether an answer is correct, and elaboration is the informational aspect of the message, providing relevant cues to guide the learner toward a correct answer" (Shute, 2008, p. 158).

Haefele (1993) writes that two general types of assessment, formative and summative, exist and both have unique feedback that accompany them. Formative teacher evaluation "is concerned with the development and improvement of teacher performance" (Haefele, 1993, p. 21). On the other hand, summative teacher evaluation is more relevant to and concerned with hiring, termination, and promotion decisions (Haefele, 1993). Ovando (1992) adds a third type to the list offered by Haefele (1993): "Diagnostic: To determine the presence, or absence, of knowledge, experiences, skills, and values" (p. 3).

Furthering the specificity of the definition of feedback, Ovando (2005) provides a seven-step process for giving feedback:

1. Jointly, set a climate of respect and trust.

2. Know the teacher as a person and as a professional.

32

3. Collaboratively, clarify expectations for performance.

4. Collect pertinent classroom performance data.

5. Analyze and reflect upon collected data.

6. Deliver the feedback acknowledging strengths, identifying areas of reflection and providing suggestions.

7. Follow-on (providing on-going support and resources) and encourage teachers to excel. (p. 173)

A decade before he laid out the feedback process in seven-steps, Ovando (1992) listed eight characteristics that feedback must have:

Relevant: Addresses student and teacher specific achievements, needs, and interests as well as specific learning and teaching behaviors.

Immediate: Provided as soon as information about student and teacher performance is available.

Factual: Based on actual student achievement (performance on a test, assignment, or project) and teacher's instructional behaviors.

Helpful: Provides suggestions for improvement of teaching and learning.

Confidential: Given directly to student or teacher without an intermediary.

Respectful: of student's and teacher's integrity and needs.

Tailored: Designed to meet individual student or teacher's specific needs and circumstances.

Encouraging: Motivates student and teacher to continue and to increase teaching and learning efforts. (p. 5)

While Ovando's list of requirements may appear comprehensive, his is hardly the lone list of mandates for feedback. Another list stipulates that feedback should:

1. Descriptive rather than evaluative

2. Specific rather than general

3. Takes into account the needs of both the receiver; and the giver of feedback.

4. Directed toward behavior that the receiver can control.

5. Solicited rather than imposed.

6. Well-times (i.e. immediate).

7. Checked to insure clear communication. (Fedor & Buckley, 1987, p. 172)

While both lists lay out important criteria for giving feedback, it is important to understand that feedback is not one-dimensional. There is more than one type of feedback. Hattie and Timperly (2007) write that "there is a distinction between feedback about the task, about the processing of the task, about the self-regulation, and about the self as a person" (p. 89). Of these, Hattie and Timperly (2007) assert that feedback about the task is the most powerful, while feedback at the personal level is rarely effective. Furthermore, feedback is only effective "to the extent that it increases the recipient's perceptions of competence" (Fedor & Buckley, 1987, p. 174).

In an attempt to arrive at an operational understanding of feedback, DeNisi & Kluger (2000) created what they called feedback intervention theory which is based on five basic assumptions:

1. Behavior is regulated by a comparison of feedback with a goal or standard.

2. Goals or standards are arranged hierarchically.

3. Attention is limited, so only those feedback-standard gaps that receive attention will regulate behavior.

4. Attention is normally directed to a moderate level in the hierarchy.

5. Feedback interventions change the locus of attention and so affect behavior.

(p. 131)

DeNisi & Kluger (2000) go on to state that "there is more to feedback intervention theory, but the major point is that the effectiveness of any feedback intervention depends on the level at which the intervention focuses our attention" (p. 132). They also echo previously discussed scholars when they assert that feedback should "focus on the task and task performance only, not on the person or any part of the person's self-concept" (p. 134).

All of these disparate threads of meaning of and requirements for feedback can be pulled together as—simply put—information communicated to learner in order to close some knowledge or skill gap the learner possesses (Sadler, 1989). In other words, feedback is a tool for growth.

Prior to moving forward in this discussion of feedback, it would be prudent to pause and note that the overall coconscious on the definition of feedback is accompanied by an equally strong overall note of caution regarding the findings contained in the body of feedback research and resulting literature. DeNisi & Kluger (2000) note that "everyone is interested in performance feedback—knowing how well he or she is performing some task" (p. 129). Despite this universal desire, feedback remains a somewhat tenuous thing:

Despite the plethora of research on the topic, the specific mechanisms relating

feedback to learning are still mostly murky, with very few (if any) general

conclusions. Researchers who have tackled the tough task of performing meta-

analyses on the feedback data use descriptors such as "inconsistent,"

"contradictory," and "highly variable" to describe the body of feedback findings.

Ten years later those descriptors still apply. (Shute, 2008, p. 156)

Further mudding the waters, as noted by Balcazar, Hopkins, and Suarez (1985), attempts

to quantify and understand the effectiveness of feedback are complicated by the many

different characteristics of feedback: "…feedback may be differently effective depending

on its characteristics…. Unfortunately, very little research has identified the differential

impact on performance produced by the various characteristics of feedback" (p. 66).

Along the same blurry lines, Fedor and Buckley (1987) note that the frequency of the

feedback is important (p. 172); however, rather than providing a point of clarity in the

understanding of feedback, they go on to note that "…the issue of how often feedback

should be provided to organizational members has not been a main focus in the feedback

literature and research" (p. 175). Here the authors note that the frequency of the feedback

is important and then point-out that this claim has not been a primary focus of the

relevant literature and research.

Shute (2008) provides the most succinct indictment of the body of knowledge

regarding feedback: "Within this large body of feedback research, there are many

conflicting findings and no consistent pattern of results" (p. 153). And, finally, "despite

its central impact on learning, feedback is still relatively underexplored, and is a process

which faces challenges, such as time, miscommunication and emotional barriers" (Carless, 2006, p. 220).

With a definition for and limitations of feedback, as codified in the literature, established the next important aspects of feedback to be discussed here are the need for and use of feedback as presented in the relevant literature. The discussion of the need for and use of feedback as presented in the relevant will be addressed in three segments: teachers appreciate it, feedback is a means of improving teacher ability, and relevant study findings.

Perhaps the simple concept of teachers appreciating feedback warranting it is, well, lacking in depth, it is still worth noting. Ovando (1992) writes that "teachers also see feedback as a means to acknowledge effective teaching practices, to identify areas of need, and to provide suggestions for improvement" (p. 4). A decade later, referencing a study in which teachers were provided prompt feedback, Ovando (2005) noted that "according to the participants, all teachers responded in a positive way. Teachers were pleased with the level of specificity of the feedback" (p. 177). The appreciation of feedback, and specific feedback in particular, is worth noting as the quality of the feedback provided to teachers by the administrators who supervise them is central to the purpose of this study. Clearly, teachers show a preference for timely specific feedback, this is established here and in the definition conversation above; however, as will be discussed in chapter four, what is preferred, appreciated, or desired is not always what is received.

Among the many aspects of the definition of feedback discussed above, the recurring concept is that feedback ought to improve a learner's ability to perform a particular task. In this study that task is teaching and the learner(s) are teachers. Those providing the feedback—doing the guiding & teaching—are the school-based administrators, the instructional leaders: "As an administrator, your 'students' are your teachers. The whole school is your classroom, and you need to treat it like your classroom" (Jerald, 2012, p. 15). The importance of the principal in the feedback-to-improve-teacher-performance process is noted in the literature:

> Principals can positively influence classroom instructional practices that in turn improve student learning. Moreover, principals also can be instrumental in creating working conditions and opportunities that encourage and allow teachers to learn from and with one another to improve student learning. Because of that combination of direct and indirect impacts, the researchers concluded, "principals are the most important actor in student learning." (Jerald, 2012, p. 9)

In fact, the research shows that "all available evidence suggests that feedback interventions often have exactly the effect they are intended to have—they help employees to improve their performance" (DeNisi & Kluger, 2000, p. 130). This concept of feedback leading to worker-improvement is supported in the education context by studies (Ovando, 2005). Furthermore, the literature shows that "feedback which is constructive and meaningful may lead to successful teaching and learning, as well as to personal satisfaction" (Ovando, 1992, p. 6). Not only is it shown in the literature that feedback may lead to successful teaching, it is asserted that "changing classroom

instruction to affect students' academic achievement requires that teachers benefit from constructive feedback…" (Ovando, 2005, p. 179). Written feedback in particular has been shown to have potential to "guide teacher's professional development aimed at enhancing teaching and learning" (Ovando, 2005, p. 177). Finally, it has been concluded that "feedback increased both learning and motivation" (DeNisi & Kluger, 2000, p. 130).

It is worth noting here that—despite the noted positive attributes of feedback noted in the literature—the essential nature of feedback to teacher improvement is not the key take-away some leave the relevant literature with. Indeed, modern education policy—at the federal level—has been shaped by other important findings noted in the literature:

> Motivated in part by research in economics on the importance of teachers in education production, the federal government has encouraged states and school districts to use student achievement growth to measure teacher effectiveness as part of the incentives built into its $4.3 billion Race to the Top Fund. Encouraged by federal incentives, many states are moving quickly to establish new regulations for how teachers are evaluated and compensated, and researchers will have great opportunities to study the impact of these new measures and new programs on productivity in public education. (Kane, Staiger, Rockoff, & Taylor, 2012, p. 3210)

Despite this federal emphasis on teacher evaluation, it should be noted that the bureaucratic process of teacher evaluation and feedback is fraught with problems, rarely results in "recognition of a teacher or the termination of his or her employment, the

improvement of curriculum or program activities, or the deployment of staff development resources to meet teachers' specific instructional needs (Darling-Hammond & Wise, 1984, p. 29). Additionally, teacher evaluation "cannot differentiate levels of competence, nor can it produce values rewards or useful advice" (Darling-Hammond & Wise, 1984, p. 33).

The cautionary notes provided by Darling-Hammond & Wise (1984) notwithstanding, the overall body of contemporary research is more favorable regarding feedback and evaluation, especially written feedback (Ovando, 2005). The importance of the competence of the instructional leader delivering the feedback is noted in the literature as is the fact that feedback from those in leadership positions is "more frequently associated with consistent effects than feedback from any other source…" (Ovando, 2005; Balcazar, Hopkins, & Suarez, 1985, p. 81).

The emphasis of the discussion thus, and indeed the focus of this study, is on written feedback, it is worth noting that other forms of feedback are also noted alongside written feedback as being of benefit to learners. In reference to a study of verbal praise and feedback, Cameron and Pierce (1994) note that:

> The present findings suggest that verbal praise and positive feedback enhance people's intrinsic interest. This is an important finding. Most social interaction in business, education, and clinical settings involves verbal feedback from managers, teachers, and therapists. When praise and other forms of positive feedback are given and later removed, people continue to show intrinsic interest in their work. (p. 397)

This shows that although this particular study, and much of the literature, are concerned with written feedback, verbal is also a route to access the benefits of feedback.

Despite the well-documented potential-positive outcomes of feedback discussed above, the results have not consistently matched what the literature would indicate should be happening. An example of feedback proving the research and accompanying body of literature correct can be found in the well-established Teacher Evaluation System (TES) used by the Cincinnati Public Schools. One comprehensive study of TES found that:

> Students assigned to a teacher undergoing TES evaluation score 0.072 standard deviations higher in math, on average, than students assigned to the same teacher in years before she participated in TES. And students assigned to teachers in years after the teacher participates in TES score 0.111 standard deviations higher in math on average. In other words, we would expect students' test scores to be higher if their teacher has participated in TES. (Tyler & Taylor, 2011, p. 19).

It is exceedingly important to note here that the TES is built on a foundation of specific, written feedback that is based on Charlotte Danielson's Enhancing Professional Practice: A Framework for Teaching and provided to the teacher by an observing administrator (Tyler & Taylor, 2011, p. 4). The specific, written feedback component of TES is crucially important as it is part of why TES works and, in other circumstances, feedback fails to function at its fullest potential. One study of feedback conducted in a Florida school district (Rafalski, 2015) offers an insight into why one state—Florida—is not seeing the teacher-improvement connected to feedback that the study of TES showed.

Rafalski (2015) found the quality of the feedback provided to the teachers by administrators to be lacking the specificity present in TES:

> The findings from this research study showed that the majority of the feedback
> provided was a retelling of the classroom events during a given observation rather
> than feedback that was informative, constructive, objective, actionable, and
> focused on specific classroom strategies and behaviors during a set time interval.
> Even the no feedback level, which meant the comment section was left blank, had
> a higher percentage than did standardized feedback and reflective questions
> combined. (p. 135)

Rafalski (2015) shows that although feedback does have well-documented positive potential, the simple act of submitting any verbiage in a box labelled "feedback" does not automatically equate to actual feedback, and, as such, does not by the mere fact of it occurring enjoy the well-documented positive benefits of feedback. Looking at Rafalski (2015) and the study of TES (Tyler & Taylor, 2011) together it can be seen that quality—specific and targeted—written feedback can result in teacher improvement that can be observed and measured by student academic achievement; however, feedback that fails to meet the level of specificity and targeted nature of that offered as part of TES does not result in the same caliber of improvement. The mere act of feedback—absent quality of feedback content—does not guarantee results: The quality of the feedback is relevant.

The final lens through which the discussion of feedback in the relevant literature will be viewed through is that of qualifying commentary made regarding feedback, especially coverage in the literature of the role the administrator plays in the feedback

process. An example of qualifying commentary would be the requirements around the suggested—for best results—frequency of providing feedback: "The best bet is that feedback should be provided at least weekly or biweekly" (Balcazar, Hopkins, and Suarez, 1985, p. 84). Here and moving forward, feedback can be understood as having the main purpose of reducing "discrepancies between current understandings and performance and a goal" (Hattie & Timperly, 2007, p. 86). Furthermore, it should be noted that "to take on this instructional purpose, feedback needs to provide information specifically relating to the task or process of learning that fills a gap between what is understood and what is aimed to be understood" (Hattie & Timperly, 2007, p. 82). Prior to moving forward to a discussion of the crucial role administrators—assistant principals, instructional coaches, and principals—play in the feedback process, it should be, again, noted that feedback is note without its failings and limitations: "The results [of a meta-analysis of feedback data] indicated that, although feedback interventions were usually effective, in more than one-third of the cases feedback actually lowered subsequent performance" (DeNisi & Kluger, 2000, p. 129).

Now, to conclude the discussion of the coverage of feedback in the relevant literature, the specific qualifier that is the feedback-giving administrator and the role her or she plays in the feedback process, as covered in the literature, will be discussed. It has been noted in the literature that "those performing instructional supervision functions need to enhance their capacity and ability to deliver accurate feedback as well as to reflect and practice their feedback delivery skills" (Ovando, 2005, p. 174). Additionally, as noted by Kane & Staiger (2012) study, "high-quality classroom observations will

require clear standards, certified raters, and multiple observations per teacher" (p. 29). Not only has the literature established the importance of the administrator's abilities, it has left little ambiguity about the influence an administrator plays on student performance: "…there is mounting evidence that principals influence student achievement in their role as organizational and instructional leaders" (Kane, Staiger, Rockoff, & Taylor, 2012, p. 3185). In order to effectively execute their role as instructional leaders and harness the power of feedback to positively influence student achievement outcomes on their campus, principals need to master two kinds of expertise (Jerald, 2012, p. 11). Writing about these two types of expertise principals must master to be effective instructional leaders, Jerald (2012) states:

> They [principals] need to be experts in classroom instruction who are able to observe and analyze teaching practices to pinpoint precise areas of strength, identify opportunities for improvement, and pose questions for further inquiry. And they need to be experts in adult learning who can plan and deliver the kinds of supports teachers need to acquire new knowledge and skills. Principals must rely on both sets of expertise to gather deep and detailed evidence about the quality of instruction in all of their classrooms, based on which they can work with other instructional leaders to provide targeted supports for teaching improvement. (p. 11)

This echoes the reality that, regardless of the industry, the goal of any employer when evaluating an employee is to improve employee performance (Tyler & Taylor, 2011, p. 1). The goal of the principal is to improve student achievement outcomes by improving

teacher instructional ability through evaluation and feedback. Despite this fact, only 8 to 17 percent of a principal's day is spent on "activities related to instructional leadership, and some evidence suggests that half of those activities lack sufficient focus to have any real chance of helping teachers improve instruction" (Jerald, 2012, p. 2). Donaldson (2010) provides an illustration of these activities lacking sufficient focus to have any real chance of helping teachers with the example of Patricia Hopkins. Hopkins, at the time the newly appointed superintendent of two school districts in north Maine, found much of the evaluative feedback provided to educators in her districts to be meaningless, vague, platitudinal, banal, and otherwise useless to the educators receiving it: "As she read through the evaluations, she found that many were full of 'valentines'—her word for vague, meaningless praise—and largely devoid of constructive criticism or concrete feedback" (p. 54). Perhaps due in no small part to this documented phenomenon of poor feedback and the importance of feedback, "a growing number of school systems are requiring candidates [for assistant principal and principal positions] to demonstrate that they can accurately observe lessons and provide feedback to teachers or use data on teacher and student performance to plan and deliver high-quality professional development for teachers (Jerald, 2012, p. 6). Indeed, because "individuals, like organizations, depend upon the information they receive from their environment in order to make adjustments in their work activities" (Fedor & Buckley, 1987, p. 171), the days of district-level leaders wanting principals to be better instructional leaders are gone; now principals must be instructional leaders who partner with teachers and other academic support staff to ensure high-quality instruction on their campus (Jerald, 2012).

45

The Marzano Evaluation Model

For the 2015-16 school year—the year from which feedback was collected and analyzed for this study—the school district of Osceola County made use of the Marzano evaluation model. Given the importance of the Marzano evaluation model in the process of producing the feedback that was analyzed for this study, it seems warranted here to spend some time reviewing the Marzano evaluation model and the literature relevant to that particular model.

The Marzano evaluation model is one among several teacher-evaluation models available for adoption by school districts in the state of Florida:

The Marzano Evaluation Model is currently being used by the Florida Department of Education (DOE) as a model that districts can use or adapt as their evaluation model. That Marzano Evaluation Model is based on a number of previous related works that include: *What Works in Schools* (Marzano, 2003), *Classroom Instruction that Works* (Marzano, Pickering, & Pollock, 2001), *Classroom Management that Works* (Marzano, Pickering, & Marzano, 2003), *Classroom Assessment and Grading that Work* (Marzano, 2006), *The Art and Science of Teaching* (Marzano, 2007), *Effective Supervision: Supporting the Art and Science of Teaching* (Marzano, Frontier, & Livingston, 2011). Each of these works was generated from a synthesis of the research and theory. Thus the mode can be considered an aggregation of the research on those elements that have traditionally been shown to correlate with student academic achievement. (Marzano, 2011)

46

Of the 67 counties in Florida currently making use of a teacher evaluation system, 25 of them employ the Marzano model (Rafalski, 2015, p. 69). According to Phillips (2014), "The Marzano causal teacher evaluation system is a performance appraisal system. It requires student performance outcomes to be combined with observable and measurable teacher behaviors to determine a teacher instructional practice score" (p. 25). The Marzano Evaluation Model is based on the understanding gained from analyzing the data from "thousands of studies that span multiple decades" (Marzano, 2011, p. 5).

The Marzano Evaluation Model is composed of four domains: classroom strategies and behaviors, preparing and planning, reflecting on teaching, and collegiality and professionalism (Marzano, 2011, p. 1). The four domains are composed of 60 elements: "41 in Domain 1, 8 elements in Domain 2, 5 elements in Domain 3 and 6 elements in Domain 4" (Marzano, 2011, p. 1). While each domain is relevant to teacher performance, "of the four domains, Domain 1 has the largest and most direct impact on student learning" (Phillips, 2014, p. 26). During informal and formal observations by administrators, teachers' use of the 41 elements in Domain 1 are rated using a proficiency scale (Marzano, 2011). The proficiency scale rates a teacher's use of the element as either innovating (4), applying (3), developing (2), beginning (1), or not using (0) (Marzano, 2011). It is important to note that "only the elements being used or called for should be evaluated during a classroom observation" (Phillips, 2014, p. 26).

Marzano (2011) claims that "experimental/control studies have been conducted that establish more direct causal linkages with enhanced student achievement that can be made with other types of data analysis" (p. 6). Marzano (2011) also asserts that

"correlation studies (the more typical approach to examining the viability of a model) have also been conducted indicating positive correlations between the elements of the model and student mathematics and reading achievement" (p. 6).

One such correlation study was conducted in Oklahoma involving 59 schools, 117 teachers, and over 13,000 k-12 students (Marzano, 2011, p. 5). The results of the study were that "96% of the 82 correlations (i.e., 41 correlations for mathematics and 41 for reading) were found to be positive with some as high as .40 and greater. A .40 correlation translates to an effect size (i.e., standardized mean difference) of .87 which is associated with a 31 percentile point gain in student achievement" (Marzano, 2011, p. 6). Furthermore, Marzano (2011) states that "these studies also aggregated data across the nine questions in Domain 1. All correlations were positive for this aggregated data. Seven of those correlations ranged from .33 to .40. These correlations translate into effect sizes of .70 and higher" (p. 6). Additionally, as previously noted, the effectiveness of the Marzano evaluation model has been established not only through correlation studies, but also through experimental/control studies:

> To date over 300 experimental/control studies have been conducted. Those studies involved over 14,000 students, 300 teachers, across 38 schools in 14 districts. The average effect size for strategies addressed in the studies was .42 with some studies reporting effect sizes of 2.00 and higher. An average effect size of .42 is associated with a 16 percentile point gain in student achievement. (Marzano, 2011, p. 5)

Research (Ashley, 2015) has also shown that proper training improves the ability of administrators to use the Marzano model, which, in turn, improves student achievement outcomes by maximizing the desirable effects of the Marzano evaluation model.

Implementation of or use of the Marzano evaluation model is outlined in *The Teacher Development Tool Kit* (Livingston & Livingston, 2012). According to the prescriptions in Livingston & Livingston (2012), beginning or initial status or experienced or professional status are the two groups that teachers should be divided into for observation purposes. Again, according to the process delineated in Livingston & Livingston (2012), beginning teachers should have two formal observations by the middle of the school year and two more by the end of the school year; meanwhile, experienced teachers should have only one formal by the middle of the school year and one additional formal by the end of the school year. The observation process is not limited to, nor fully engaged in by the mere act of an administrator sitting in a teacher's classroom:

> The process for formal observations includes an administrator conducting both pre- and post-conferences with teachers. The pre-conference is conducted prior to the observation to gather information about the observation, identify key elements to be observed, and to review the teacher's lesson. During the post conference, the supervisor ratings are discussed and student evidence may be reviewed. Teacher evaluation ratings can be adjusted after the post-conference. (Phillips, 2014, p. 27)

Prior to participating in the pre-conference, the teacher should identify the specific instructional element(s) they intend to focus on for improvement during the upcoming or

current school year (Marzano, 2011). This act of selecting an instructional element or instructional elements to focus on for improvement is often referred to as the deliberate practice. Ericsson, Krampe, and Tesch-Romer (1993) define deliberate practice as the process by which professionals refine and improve their practice and skills. It is important to note that while this particular study focuses on the written feedback resulting from formal observations, the literature also recommends classroom walkthroughs—brief class visits during which a supervisor observes only a few instructional elements—as a component of best practices in using the Marzano evaluation model (Livingston & Livingston, 2012).

Finally, as noted by Marzano (2013), "Given that forty-one of the sixty elements in the model are from Domain 1, the clear emphasis in the Marzano model is what occurs in the classroom" (p. 2).

Communication

While The Lake Wobegon Effect provides one possible explanation for the phenomenon of teacher rating inflation and feedback along with the Marzano Evaluation Model are important components of the teacher-evaluation process, communication—at the most foundational level—is the basis for and an essential component to understand in the teacher-evaluation and feedback process. Although much of the literature dealing with communication is outside the scope of this study, a cursory discussion of the relevant literature dealing with communication will be provided here.

In 1924 researchers went to work at the Hawthorne Works of Western Electric in Cicero, near Chicago; their goal was to determine the optimal illumination level for maximum worker productivity (Owens & Valesky, 2015, p. 124). While they failed to discover an optimal level of illumination, what they did find has had a lasting impact on how researchers think about work, workers, and supervisor-to-worker communication: "The workers in the experimental group were responding to their perceptions of the expectations of the experimenters and not to the changes in the physical environment. Thus, the workers were responding to psychological factors…" (Owens & Valesky, 2015, p. 125). The fact that there is a relationship between worker productivity, psychological factors, and communication is now often referred to as the "Hawthorne Effect" (Owens & Valesky, 2015). What the Hawthorne Effect established is the power of perceived supervisor expectations to positively influence worker productivity. Given that even an expectation merely perceived can improve worker productivity, it seems a safe certainty that overtly communicated supervisor-to-worker expectations will improve worker productivity.

In order to discuss supervisor-to-worker communication of expectations, the basic concept of what communication is must be understood. What is communication? Dainton and Zelley (2014) highlight the complicated social-understanding of communication:

> Communication is perceived as a magical elixir, one that can ensure a happy long-term relationship and can guarantee organizational success. Clearly, popular culture holds paradoxical views about communication: It is easy to do yet powerful in its effects, simultaneously simple and magical. (p. 1)

The most crucial component of this complicated social-understanding is the paradoxical and complex nature of communication in society. Despite its ubiquity, communication is not simple.

Other understandings of communication include the basic understanding of it as the flow of information from one person to another (Axley, 1984), or simply one activity in a series of things organizations do (Deetz, 1994). Communication theorists, however, define it as the process by which people interactively create, sustain, and manage meaning (Conrad & Poole, 2012). The essential component of this definition, for the purposes of this study, is the term interactive. Communication, in this understanding, requires something from both parties involved in order to arrive at the creation, stasis, or management of meaning. To bring this back to The Hawthorne Effect, both workers and those seeking to influence their behavior by shaping their perceptions of expectations are involved in a give and take that results in meaning, or, in the case of workers, an increase or decrease in productivity.

<div align="center">Interpersonal Communication</div>

The next topic for consideration in this discussion of the relevant literature is interpersonal communication, and more specifically interpersonal behavior as a subset of interpersonal communication.

The study of interpersonal communication is a branch off of the larger tree of academic study that is communication theory. Some (Moore, 2017) would contend that the school of thought concerned with interpersonal communication came into focus in the

1960s when scholars began "extending postpositivist theories from psychology" (p. 1). Meanwhile others (Braithwaite & Baxter, 2008) argue that the true roots of the line of intellectual inquiry that concerns itself with interpersonal communication dates all the way back to ancient Greece and Rome and possibly even further back to Africa and China (p. 1). While the scholars may disagree on the exact moment the concept of interpersonal communication entered the intellectual lexicon, what is not debated is the role the post World War II era played in the development of this field of study:

> After World War II, speech teachers also began teaching courses in small group discussion. Social scientists, especially in psychology, began studying persuasion and obedience to authority, trying to understand the process of interpersonal influence to help explain some of the atrocities that happened during that war. While they were interested in the psychology of persuasion, many also realized that we needed to study how persuasion was enacted. (Braithwaite & Baxter, 2008, p. 2)

By the 1970s & 1980s interpersonal communication studies had spawned its own offshoots, including the study of nonverbal communication, gender, and intercultural communication among others (Braithwaite & Baxter, 2008). The field has continued to grow and expand to the present day.

Defining interpersonal communication is no less difficult than nailing-down the precise origin story for the field of study. Dziak (2016) provides a definition that fits nicely in the larger context of communication theory:

Interpersonal communication is the process by which people exchange

information with others. It occurs when one party sends a message and another

party receives the message. Messages may be verbal or nonverbal, depending on

whether they employ language. Interpersonal communication may involve two

people or larger groups, and may take place in many varied settings and

situations. Direct interpersonal communication involves face-to-face exchange of

information, while mediated interpersonal communication uses technological

means to transfer messages. (p. 1)

While Dziak (2016) provides a straight-forward definition that neatly fits into the

framework of previously discussed concepts of communication, other scholars

conceptualized interpersonal communication as a vastly more complex act that human

beings engage in. John Stewart (1999) provides an understanding of interpersonal

communication housed within the understanding that humans are the architects of their

own reality:

Communication is the way humans build our reality. Human worlds are not made

up of objects but of peoples' responses to objects, or their meanings. And these

meanings are negotiated in communication. Try not to think of communication as

simply a way to share ideas, because it's much more than that. It's the process

humans use to define reality itself. (p. 25)

While both of these definitions encompass the conveying of ideas, one—Stewart's—has

much grander implications. The differences present between these two definitions speak

to the vast variance in this particular field of inquiry.

In simplest terms, interpersonal behaviors can be understood as the mechanisms through which humans engage in interpersonal communication. Reduced to their basic aspects, these behaviors can be shorted into two categories: (1) assertive and (2) aggressive; assertiveness carries positive connotations and is seen as advantageous while aggressiveness carries negative connotations and is seen as undesirable (Mauger & Adkinson, 1980).

While a summary discussion of interpersonal communication theory has been provided above, this discussion alone is insufficient to provide a thoughtful understanding of interpersonal communication as relevant to this study. In order to reach a more thorough understanding of interpersonal communication as a discipline of study, six closely related theories that support, substantiate, qualify, expand, or buttress thinking around interpersonal communication in some important way will be explored in the pages that follow. These theories will be introduced, their historical context will be provided, and their relevance to interpersonal communication in general and this study in particular will then be discussed. The six related theories are: symbolic interactionism, coordinated management of meaning, interpersonal deception theory, social penetration theory, uncertainty reduction theory, relational dialectics, and the interactional view.

Although he never published his work on the theory, George Herbert Mead is considered to be the father of symbolic interactionism (LaRossa and Reitzes, 1993). Herbert Blumer, a student of Mead's, is credited with coining the term "symbolic interactionism" and putting the basic tenants of the theory down in writing (Blumer, 1969). In his writing, Blumer put forth the concept that humans interact with other people

or things based on meanings ascribed to those people or things; and meaning—which is a central aspect of human behavior—is derived from interactions with others and from society (Blumer, 1969). A simple example of this theory would would be that a teacher who—despite research to the contrary—believes homework is essential to quality education, developed that meaning for the value of homework through previous interactions—most likely an influential former teacher or parents who place a high value on homework. An administrator who did not have similar formative interactions would—in a symbolic interactionists understanding of things—arrive to a discussion of homework policies with a different meaning for homework based on his or her prior interactions.

A little over a decade after Blumer (1969) coined "symbolic interactionism", Pearce and Cronen (1980) developed coordinated management of meaning (CMM) theory. While Blumer (1969) posited that meaning is central and derived from prior interactions, Pearce and Cronen (1980 & 1982) put forth the idea that rather being derived from prior social interactions, meaning is created by individuals during conversation or interaction. CMM is a rules based theory; navigating these rules to coordinate meaning becomes difficult when the two parties involved have differing views, differing values, and/or differing goals (Pearce and Cronen, 1980). Applying these tenants from Pearce and Cronen (1980) it is easy to understand why a school administrator who values student compliance over relationship building and a teacher who holds reversed values struggle to communicate.

It is easy to understand the sincere struggles to communicate and reach a place of shared understanding that occurs between two communicators with different value system founded in different lived experiences; however, it is often more a struggle to understand miscommunication furthered by deception. People engage in deception for many different reasons (Buller and Burgoon, 1996). In the context of relationships with others, people lie to "avoid hurting or offending another another person, to emphasize their best qualities, to avoid getting into a conflict, or to speed up or slow down a relationship" (Buller and Burgonn, 1996, p. 97). Every act of deception has at least three aims—to accomplish a specific tasks, to establish or maintain a relationship with the other party, and to sustain the good image of one or both the parties involved in the deception; this includes all deceptions from well-intended lies told to spare the deceived from a harsh truth to elaborate deceptions crafted to advance or better the author of the deception (Buller and Burgoon, 1996). An administrator who rates a mediocre teacher as "highly effective" and offers only praise or benign criticism-free feedback to said teacher, is certainly engaging in deception. While the literature shows that it is certainly plausible that this deception is done to spare the feelings of the teacher, the literature also shows it is equally plausible that the theoretical framework of this study holds true: The administrator is perpetuating the ruse the teacher's exceptional abilities to mollify the teacher while maintaining the trappings of bourgeois status with little to no real conflict. The deception turns the Marxian constant state of conflict into little more than vacuous verbiage, so many banal platitudes.

Social penetration theory does not deal with the outright lies and deceit of interpersonal deception theory; however, it does deal, in its way, with omission and obfuscation. According to Altman & Taylor (1973), individuals consist of a public persona and a private self, and, as relationships between individuals develop, the public persona is penetrated in order to reach the core personality of the parties involved—this is achieved as the parties reveal things about themselves. Individuals begin by revealing low-risk information—a preference for Italian cuisine over others—and only reveal more personal information when the gesture is returned in kind (Altman & Taylor, 1973). Given that the perceived risk or cost of revealing information increases in direct relationship to how close that information lies to the core of a person, it is highly unlikely that an individual will reveal deeply, core personal information to many people (Altman & Taylor, 1973). Administrator-to-teacher feedback and teacher effectiveness ratings are certainly not the kind of deeply personal subject matter fraught with layers of risk; however, an administrator who reveals nothing of him or herself and is, at all times, revealing only a public persona—or, in this context, an administrator persona—will struggle to create communication connections that go beyond the superficial.

Social penetration theory dealt with the layers people are composed of and the degree to which we allow others to penetrate our social persona to reveal our core self. Uncertainty reduction theory deals with the uncertainty brought on by people we do not know, especially when we are made to rely on them in some fashion, and our desire to reduce this uncertainty by getting to know the other person (Berger and Calabrese, 1975). The theory is based on the predictable pattern of information gathering that occurs when

58

people interaction with someone they are not familiar with (Berger and Calabrese, 1975). Uncertainty reduction and social penetration can be understood as foundation of a tension that arises due competing desires: The desire to reduce the uncertainty brought on by a new person and the desire to not reveal too much about ourselves—knowing full-well that the new person will not reveal things to reduce our uncertainty unless we do so as well. Again, this theory speaks to the need for administrators to be humanized and known by the people they seek to lead.

While symbolic interactionism asserts that meaning is derived through interactions and society (Blumer, 1969) and coordinated management meaning asserts that individuals construct their individual meanings through and while engaging in conversation with others (Pearce and Cronen, 1980), the interactional view posits that a relationship cannot be understood through either individual party involved (Watzlawick, Weakland, and Fisch, 1974). Unlike the other theories of communication discussed here, the interactional view frames communication within the social and cultural context in which it occurs (Watzlawick, Weakland, and Fisch, 1974). According to Watzlawick and others (1974) people cannot not communicate, even sitting silently sends a message. Griffin (1997) asserts that healthy relationships are composed of both symmetrical and complementary communication. Symmetrical communication occurs when communication is based on equal power, while complementary communication occurs when differences in power exist (Watzlawick, Weakland, & Fisch 1974). Given the inherent power imbalance in administrator-to-teacher communications, there is a predisposition to not meet the criteria for healthy relationships laid out by Griffin (1997).

An administrator who sees a teacher as an opponent in a Marxian conflict—as discussed in the theoretical framework of this study—is at risk of unhealthy relationships with those he or she is charged with leading. This reason is why the interactional view's inclusion of social context—a workplace that comes loaded with hierarchy and the power imbalances that engenders—is an important theory of interpersonal communication to consider in the context of this study.

<div align="center">Conflict Avoidance</div>

The next point of emphasis in this review of the relevant literature is a look at the coverage of conflict avoidance in the literature. It is important to note here that while conflict avoidance is a human behavior that can be observed and studied in many different contexts (Mauger & Adkinson, 1980) this discussion will focus on coverage in the literature relevant to conflict avoidance in the workplace, specifically the schoolhouse as the workplace.

Before moving into what the literature has to say about conflict avoidance in the schoolhouse-workplace, it is important here to establish a foundational understanding of what conflict avoidance is. For the purposes of this discussion, conflict avoidance can be understood as fight phobic behaviors (Bach & Wyden, 1968). A person who has a tendency toward or engages in conflict avoidance would "tend to evade open disagreement or conflict with others" (Mauger & Adkinson, 1980, p. 5).

Earlier in this chapter, the Lake Wobegon Effect was offered as one possible means of understanding or explaining the phenomenon of inflated teacher evaluations.

Now, the desire—on the part of administrators—to avoid conflict, will be explored as

another means of understanding the reoccurring phenomenon of inflated teacher

evaluations. According to one study (Langlois & Colarusso, 1988) the "vast majority" of

teachers are awarded top ratings on their evaluations (p. 13). While it is plausible that

these ratings are due to unprecedented emergence in vast and widespread teacher

greatness, it is more likely that the ratings are due to a failure to fully perform their duties

on the part of administrators who supervise teachers: "The ratings indicate that school

executives often fail to observe and evaluate teachers—or that they overrate the teachers

they do evaluate" (Langlois & Colarusso, 1988, p. 13). This failure to execute an

essential function of leadership and supervision is egregious given the power and

influence supervising administrators have on the career of the teachers they supervise:

"Every principal holds in his hands the career of a significant number of teachers.

Hopefully, principals exercise wise judgments and view their role as one helping their

teachers do a better job with students" (Hain & Smith, 1996, p. 1).

It has been argued (Stronge, 1993) that unity of purpose is an essential component

of an effective school. Furthermore, it has been asserted that "a performance appraisal

system has its genesis in the broad purposes of the organization" (Castetter, 1981, p.

239). While untied purpose and effect performance appraisal systems drive high-

performing and successful schools (Stronge, 1995), conflict avoidance on the part of a

supervising administrator serves to sabotage the performance appraisal system and thus

undermines the functioning of the entire school.

The potential for corrosive conflict avoiding behavior exists in the fundamental process of what administrators due. In discussing the fundamental process of administration, Getzels & Guba (1957) write, "The process of administration deals essentially with the conduct of social behavior in a hierarchical setting. Structurally, we may conceive of administration as a series of superordinate-subordinate relationships within a social system" (p. 424). Here the superordinate role is held by the supervising administrator who is expected—by virtue of the administrator-as-instructional-leader paradigm—to provide instruction-improving feedback to the subordinate teacher; however, when the superordinate leader neglects their role in this process, the process breaks down. Getzels & Guba (1957) suggest that an individual should be both "adjusted and integrated" referring to an individual who is both performing up to the expectations of their role and filling their needs (p. 431). The problem arises when one has a real or self-perceived need to avoid potential conflict coupled with a role that requires one to engage in an activity, feedback-giving, that is inherently loaded with the potential for conflict. As Fuhr (1993) notes, "Many school administrators rank supervising marginal teachers as one their toughest challenges" (P. 1). For the conflict adverse administrator, this challenge becomes a monumental task. Amplifying the nature of the conflict adverse administrator and marginal teacher dynamic is the reality that "a teacher with an attitude problem must be confronted and dealt with at once, as negative attitude from any one staff member can quickly spread" (Fuhr, 1993. P. 1). Administrators cannot afford to neglect their supervisory duties, as doing so does not merely allow them to engage in

conflict avoiding behaviors, it allows poor teachers to remain in classrooms—

unchallenged—while students suffer (Langlois & Colarruso, 1988).

<center>Moral Leadership</center>

The final concept to be discussed in this chapter is moral leadership. Leadership is

"a complex phenomenon involving the leader, the followers, and the situation" (Hughes,

Ginnett, & Curphy, 1993, p. 41). Despite the inherent complexity in defining and

understanding leadership, Bernard Bass (1990) asserts that there is sufficient common

ground to arrive at a definition for leadership:

> Nevertheless, there is sufficient similarity among definitions to permit a rough
>
> scheme of classification. Leadership has been conceived as the focus of group
>
> processes, as a matter of personality, as a matter of inducing compliance, as the
>
> exercise of influence, as particular behaviors, as a form of persuasion, as a power
>
> relation, as an instrument to achieve goals, as an effect of interaction, as a
>
> differentiated role, as initiation of structure, and as many combinations of these
>
> definitions. (p. 38)

What is essential to take away from these two definitions is this: Leadership, in all of its

iterations, involves a leader interacting with followers in order to achieve some end. It is

important to note, as Thomas E. Cronin (1984) does, that while it is tempting to

romanticize leadership with images of Washington or Churchill, it is imperative to

remember that not all leaders are cut from the same cloth: "Leadership can be exercised

in the service of noble, liberating, enriching ends, but it can also serve to manipulate,

<center>63</center>

mislead and repress" (p. 27). This section focuses on two studies and one antidotal example and how the findings of those studies illuminate this negative capacity of leadership, or, as the studies show, the lapses in and decline of moral decision making by those in senior leadership positions—specifically in the education setting. A central idea to be mindful of in this section is that given both the importance and the complexity of leadership, the potential for poor decision making and its ramifications for the working environment of the subordinates of these decision makers are crucial in this study. Note that the intent of the researcher here is not to cast dispersions nor is it to label school leaders as aberrant, immoral, or otherwise failed human beings; rather, the intent here is to call attention to a pattern of findings that are necessary to consider in the context of this study.

Strenth (2013) sought to examine the moral decision making of elementary school principals, specifically elementary principals in high-poverty schools in an urban setting. From the 64 elementary schools that met the poverty criteria, receipt of Title I funds, 20 principals were selected to to take the *Defining Issues Test* (Strenth, 2013, p. 56-57). The *Defining Issues Test* (DIT) is based on the work of James Rest, and it "measures a subject's moral schemas and reasoning process used to determine the morally correct solution" (Strenth, 2013, p. 58). The study findings were, well, not what would be expected of a collection of school leaders: "The results revealed that all of the principals, except for one, operated in the moral development range from below institutionalized delinquents to just above senior high school students" (Strenth, 2013, p. 76). Indeed, the mean score for the study participants was six points below that of high school students

64

and only two points above that of prison inmates. On an inventory of moral decision making, these elementary school principals bested prison inmates by only two points.

The findings presented by Strenth (2013) were similar to those of Vitton and Wasonga (2009). Vitton and Wasonga (2009) found a mean score for elementary school principals slightly above that of Strenth (2013), placing the moral decision making of the principals in their study just above high school students and below that of enlisted sailors in the U.S. Navy. While besting high school students is certainly an improvement over eking out a slight win over prison inmates, it seems safe to say that neither group represent the kind of lofty pinnacles of moral decision making school leaders are expected to be.

Ling (2014) also sought to investigate the moral decision making of school leaders. While Strenth (2013) focused on the moral decision making of a particular group of school leaders—principals of high poverty elementary schools—Ling (2014) looked at school leadership as a continuum from graduate student in educational leadership to assistant principal to principal and sought to understand the arc of moral decision making along that continuum. Ling (2014) also used the DIT is the instrument for the study. Ling (2014) found that moral decision making declined along the leadership continuum from graduate student to assistant principal to principal. Graduate students showed the highest score, one comparable to holders of masters degrees, followed by assistant principals and then principals had the lowest score. The scores of assistant principals and principals were comparable to college juniors (Ling, 2014). These findings are less troubling than those of Strenth (2013), but the declining moral decision making along the leadership

continuum cannot be seen as a positive. Additionally, not to disparage college juniors, individuals in their very early twenties are hardly the pillars of moral decision making that school leaders are oft thought to be.

The cheating scandal that engulfed the public schools in Atlanta, GA serves as a prime example of the findings of these studies (Strenth, 2013 & Ling, 2014) in action. In 2009 erasure mark analysis showed unusually high numbers of wrong-to-right answer changes (Office of the governor, 2011). While what happened in Atlanta is riff with applicable tales of morally flawed decision making, the story of Dr. Jackie Boyce provides a particularly galling example of a lack of moral leadership. In 2009, Dr. Boyce was assigned to assist with testing monitoring at Perkerson Elementary School: "Several times during the 2009 testing, Dr. Boyce says he observed teacher Lashaine Blake, and others, improperly pointing out answers to students" (Office of the governor, 2011, p. 357). Dr. Boyce, twice, expressed his concerns to the principal, Dr. Mable Johnson, who paid his concerns no mind (Office of the governor, 2011). Dr. Boyce also noted his concerns on testing feedback paperwork, and submitted that paperwork directly to Executive Director, Dr. Sharon Davis-Williams:

> Dr. Davis-Williams did not ask Boyce any details about the teachers'
> inappropriate conduct, not did she tell him to report his observations to anyone
> else. Instead, Davis-Williams gave Boyce blank forms and directed him to fill
> them out again without the notations about teachers pointing out answers.
> (Office of the governor, 2011, p. 357).

The continuum of school leadership studied in Ling (2014) stops with school principals, the actions of Dr. Davis-Williams suggest that had Ling (2014) expanded his study to include those above the principal level, the decline in moral decision making observed in his study would have continued right up the leadership ranks.

Summary

This chapter provided a discussion of the literature and research findings relevant to this study. The discussion began with a look at The Lake Wobegon Effect—an inflating of the quality of everything. The discussion proceeded to an overview of the value of feedback in general and of feedback in the education context in particular. The chapter then discussed the Marzano Evaluation Model, the model being used by the administrators being examined in this study. Next, the chapter provided an overview of communication theory, interpersonal communication, and conflict avoidance. The chapter concluded with an exploration of the intersection of leadership and morality.

CHAPTER 3
METHODOLOGY

Introduction

The primary goal of this study is to investigate the research questions that relate to interpersonal behavior traits and the way in which they shape, if at all, the feedback provided to teachers by their assessing administrator, and, to what extent, interpersonal behavior traits can explain the noted phenomenon of inflated teacher evaluation scores as stated in chapter one. An instrument, The Interpersonal Behavior Survey (IBS), a survey that is used to develop a comprehensive interpersonal behavior profile, was used to develop interpersonal behavioral profiles for teacher-assessing administrators, the profiles were used to investigate the aforementioned research questions. The methodology employed to explore the research questions is presented in this chapter. The chapter is organized into five distinct sections: (1) research questions, (2) selection of participants, (3) instrumentation, (4) data collection, and (5) data analysis.

Research Questions

The research questions were developed using the behavior trait groupings of the IBS as model. Each question is built around the traits measured by a specific grouping presented on the IBS. The following research questions and hypotheses will guide the investigation of this study:

1. What, if any, relationship exists between the assertiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

   H_{O1}: There will be no observed relationship between assertiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise.

2. What, if any, relationship exists between the aggressiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

   H_{O2}: There will be no observed relationship between the aggressiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise.

3. What, if any, relationship exists between the relationship scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

   H_{O3}: There will be no observed relationship between the relationship scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise.

4. What is the relationship between scores above 70T on the Denial, Infrequency, and Impression Management scales—all scores that render the aggressiveness, assertiveness, and relationship portions of the respondent's survey unreliable— and the feedback provided by the administrators with those scores?

H$_{O6}$: There will be no relationship between scores above 70T on the Denial,

Infrequency, and Impression Management scales and the feedback provided by

the administrators with those scores.

<u>Selection of Participants</u>

In order to achieve the most comprehensive investigation of the aforementioned

research questions, all 131 school-based, teacher-assessing administrators in Osceola

County were invited to take the IBS and have an interpersonal behavior profile

developed. 54 school-based administrators opted to participate in this study. A

breakdown of these participants by gender, age, and years of experience in administration

can be found in chapter 1 of this study.

<u>Instrumentation</u>

Interpersonal Behavior Survey

The Interpersonal Behavior Survey (IBS) is a gender-normed 272 item, true or

false, survey—developed by Paul A. Mauger and David R. Adkinson (1980)—used to

develop an interpersonal behavior profile along four scale groupings—validity,

aggressiveness, assertiveness, and relationship—each scale grouping consists of three to

eight scales. The scales in the validity group serve to measure test-taking attitudes, and,

more importantly, these scales indicate whether a profile can be reliably interpreted. The

validity scale group consists of the denial, infrequency, and impression management

scales. The denial scale measures the respondent's hesitancy to admit to common but

socially undesirable weaknesses and feelings; including such behaviors as making fun of

others, swearing, and procrastinating (Mauger & Adkinson, 1980, p. 2). The infrequency

scale measures the respondent's tendency to "endorse items that less than 10% of the

normative sample endorsed. Such items were selected solely on the basis of frequency of

endorsement, so that the scale spans a variety of content areas" (Mauger & Adkinson,

1980, p. 2). The final scale in the validity group is the impression management scale; the

impression management scale provides a measure of the degree to which the respondent

is trying manage the impression they make with their responses. According to Mauger &

Adkinson (1980), "This scale was derived by asking college students to take the IBS

twice. On the second administration they were asked to respond in such a manner as to

create a favorable impression of themselves. Items that changed significantly comprise

the impression management scale" (p. 2).

  The second group of scales on the IBS, aggressiveness, consists of seven scales:

general aggressiveness, hostile stance, expression of anger, disregard for rights, verbal

aggressiveness, physical aggressiveness, and passive aggressiveness. The general

aggressiveness scale provides a measure of the respondent's aggressive behaviors,

feelings, and attitudes (Mauger & Adkinson, 1980, p. 4). The hostile stance scale

provides a measure of the degree to which the respondent holds an antagonistic

orientation towards other people, "a view of the world that justifies aggression in order to

get ahead in life or to protect oneself" (Mauger & Adkinson, 1980, p. 4). The expression

of anger scale provides a measure that servers as an indicator of the respondent's

likeliness to lose his or her temper and to express their anger (Mauger & Adkinson, 1980,

p. 4). The disregard for rights scale provides a measure of of the degree to which the

respondent tends to ignore the rights of others in order to protect or advance him or herself (Mauger & Adkinson, 1980, p. 4). The verbal aggressiveness scale provides a measure of the degree to which the respondent is likely to "use words as weapons" by verbally tormenting others (Mauger & Adkinson, 1980, p. 4). The physical aggressiveness scale provides a measure of the respondent's tendency to engaged in and/or his or her comfort using physical force or violence (Mauger & Adkinson, 1980, p. 4). The final aggressiveness scale is the passive aggressiveness scale. This scale provides a measure of the respondent's tendency to engage in such behaviors as stubbornness, negativism, procrastination, and complaining (Mauger & Adkinson, 1980, p. 4).

The next set of scales fall into the assertiveness group; this grouping includes the general assertiveness scale, the self-confidence scale, the initiating assertiveness scale, the defending assertiveness scale, the frankness scale, the praise scale, the requesting help scale, and the refusing demands scale. The general assertiveness scale provides a measure of the respondent's likeliness to engage in a wide array of assertive behaviors (Mauger & Adkinson, 1980, p. 4). The self-confidence scale provides a measure of the respondent's positive attitudes about him or herself (Mauger & Adkinson, 1980, p. 4). The initiating assertiveness scale "is an indication of leadership potential and the tendency to take an ascendant role in groups" (Mauger & Adkinson, 1980, p. 5). The defending assertiveness scale provides a measure of the degree to which the respondent is likely to stand up for others (Mauger & Adkinson, 1980, p. 5). The frankness scale shows a measure of the degree to which the respondent can be considered willing to clearly communicate his or her honest opinion on any given matter (Mauger & Adkinson, 1980, p. 5). The praise

scale displays a measure of the respondent's level of comfort with giving and receiving praise (Mauger & Adkinson, 1980, p. 5). The requesting help scale provides a measure of how willing and likely the respondent is to ask for reasonable favors and help when it is needed (Mauger & Adkinson, 1980, p. 5). The final scale in the assertiveness group is the refusing demands scale. This scale gives a measure of the degree to which the respondent is likely to say "no" to unreasonable or inconvenient demands (Mauger & Adkinson, 1980, p. 5).

The final group of scales, the relationship scales, consists of three scales: conflict avoidance, dependency, and shyness. The conflict avoidance scale provides a measure of the respondent's tendency to successfully or attempt to evade disagreement or conflict (Mauger & Adkinson, 1980, p. 5). The dependency scale provides an indication of the degree to which the respondent is likely to depend on others (Mauger & Adkinson, 1980, p. 5). The final relationship scale, shyness, provides a measure of the respondent's degree of introversion (Mauger & Adkinson, 1980, p. 5).

The reliability of the IBS was established "using a test-retest format over both a 2-day and a 10-week period and the coefficient alpha internal consistency procedure" (Mauger & Adkinson, 1980, p. 12). According to Mauger & Adkinson (1980) the IBS was found to be as reliable or more reliable than other personality inventories in common use (p. 12). Mauger & Adkinson (1980) report a reliability value for both 2-day and 10-week period of greater than .90 (p. 12).

Regarding the validity of the IBS, Mauger and Adkinson (1980) state that "the IBS has been correlated with several well-known personality inventories using samples

from a number of populations" (p. 17). Additionally, Mauger and Adkinson (1980) note that "convergent and discriminant validities have been assessed by noting correlations of the IBS scales with scales from these other inventories, paying particular attention to predicted relationships between scales. The General Assertiveness, Rational (SGR) scale correlated .47 with the Dominance scale of the California Psychological Inventory" (p. 17). Furthermore, the SGR correlated .63 with the Dominance scale of the Edwards Personal Preference Schedule (EPPS); it correlated .64 with the College Self-Expression Scale; it correlated .74 with the Rathus Assertiveness Schedule; and it correlated .45 with the Assertion score of the Conflict Resolution Inventory (Mauger & Adkinson, 1980, p. 17). The General Aggressiveness, Rational (GGR) scale correlated .57 with the Aggression scale of the EPPS; it correlated .47 with the Aggressive scale on the Interpersonal Check List (ICl); it correlated .55 with the Skeptical scale of the ICl and .55 with the ICL's Factor Hostility Scale. The GGR also correlated .65 with the Total Hostility Scale of the Buss-Durkee Hostility Inventory (Mauger & Adkinson, 1980, p. 17).

Interpreting an IBS profile should always begin with the validity scales. Beginning with the denial scale, any score between 60T and 70T calls into question the validity of the response; the closer the score is to 70T, the more cautious the interpreter should be. Any denial scale score over 70T means the response is invalid and a reliable full profile cannot be developed (Mauger & Adkinson, 1980, p. 20). The denial scale provides a measure of a person's "hesitancy to admit to common but socially undesirable weaknesses and feelings" (Mauger & Adkinson, 1980, p. 2). An infrequency score above

70T should also give the interpreter reason to question the response, as an infrequency score above 70T may indicate a deviant response set, possibly the result of: reading difficulties, confusion due to psychopathology, malingering or faking bad, and all-true or all-false response, or incorrect recording of responses on the answer sheet (Mauger & Adkinson, 1980, p. 21). The infrequency scale provides a measure of the "tendency to endorse items that less than 10% of the normative sample ensorsed" (Mauger & Adkinson, 1980, p. 2). Finally, a score above 70T on the impression management scale also serves to preclude the development of a reliable profile, as it shows an overt attempt to give "the right answer" as opposed to an honest response (Mauger & Adkinson, 1980, p. 21). Once the validity scales have been assessed, the interpreter can then develop a full profile based on the responses. When analyzing the responses, the interpreter should look for areas in which the respondent falls outside the normal range for a person of their gender (Mauger & Adkinson, 1980, p. 21).

Feedback Rubric

Designed by Rafalski (2015), the feedback rubric scores feedback given to teachers by administrators on a seven-level scale that measures the quality of feedback provided by evaluators to those being evaluated. Studies (Rafalski, 2015 and Butler, 2017) have employed the rubric as a means of measuring feedback.

The rubric rates feedback on seven levels. Level one on the rubric is no feedback. Level 2 is for "unrelated feedback or general statement provided" (Rafalski, 2015, p. 199). Level three is marked by a "recount of classroom events and a justification for rating", and level three includes the further clarification that "recap has several different

components (sometimes statement of percent of students being monitored or desired effect). Sometimes actually gives examples of what is wrong with no suggestion. You did this or that, teacher did this or this or that, I observed this or that..." (Rafalski, 2015, p. 199-200). Level 4 on the feedback rubric is marked by general affirmative statements, and the rubric provides the following example verbiage: "Good job, great job, excellent job, I liked, I loved, WOW!" (Rafalski, 2015, p. 199). Level 5 on the feedback rubric is a reflective question, the administrator asks the teacher a question that requires some reflection (Rafalski, 2015, p. 199). At level 6 on the rubric standardized feedback is provided (Rafalski, 2015, p. 199). The feedback rubric provides the following verbiage as example of standardized feedback: "How might you adapt and create new strategies for chunking content into digestible bites that address unique student needs and situations?"; "How might you expand your monitoring to involve more students"; "What are you learning about your students as you adapt and create new strategies?"; and "In addition to monitoring students by the use of choral responses, how else can you monitor students when chunking information?" (Rafalski, 2015, p. 199-200). The highest rating on the rubric, level 7, is marked by specific targeted feedback. Again, the rubric provides example verbiage: "1. Reference to Resource Library or Reflective Teacher 2. Maybe try.... Or You might want to try.... 3. Consider.... 4. Recommendation.... Or I would recommend.... 5. Suggestion.... Or I suggest.... 6. It might be a good idea... 7. You should.... 8. This would have been good or great if... 9. To move to a higher level, do this_____. 10. Think about..... 11. I want you to.... 12. Next time.... 13. Always.... 14. Be sure to... or Make sure you..... 15. Doing this_____would have

76

been more effective. 16. You need to.... 17. Continue to... 18. Remember this_____.

19. Do this_____. 20. This_____is a good strategy. 21. I would like to

see.... 22. Coaching idea..." (Rafalski, 2015, p. 199-200). Rafalski's feedback rubric was

developed for her study that was designed to "(a) to determine the relationship between

the number of classroom observations and teacher VAM scores and (b) to identify the

relationship between the types of feedback provided to teachers and student achievement

outcomes as measured by VAM scores" (Rafalski, 2015, p. iii).

According to Rafalski (2015) the reliability of the rubric was established using an

intercoder method:

> In order to ensure that the codes provided for comments operated reliably, an
>
> additional coder was employed to measure intercoder reliability. This additional
>
> coder received the rubric and sample of 650 comments in order to ensure that the
>
> comment definitions were followed in coding. No scored elements where
>
> evaluators left no comments were contained in the sample to determine intercoder
>
> reliability. Of the 650 sample elements with comments, 604 (92.9%) were coded
>
> identically by both coders. This suggested that the coding executed by the
>
> researcher measured replicable findings that were clear in the comments provided
>
> by evaluators. (p. 105)

It should be noted here that level 6, standardized feedback, was not employed in this

study as it refers to feedback available for evaluators to copy and paste from a platform

that was not in use in Osceola County during the timeframe the feedback examined in

this study was created. Additionally, when the researcher encountered comments that

seemed to contain elements of more than one level on the scoring rubric, the rating given always erred on the side of the higher rating. Table 6 below provides examples of administrator-to-teacher feedback scored in this study and the feedback rubric score given to that piece of feedback. Note that the feedback displayed on table 6 was selected at random to provide examples of feedback at each level of the rating rubric. Additional examples of feedback scored for this study can be found in the appendixes of this study.

Table 6. Example administrator-to-teacher feedback & accompanying rubric score.

Feedback Text	Feedback rubric score
How silly would it be for me to critique you.... You are a master teacher and I am an awe of all that you do for this school and community.	4
Students utilized their performance coach books to reinforce the concepts of the standard. Various students then volunteered to show the strategy they used to solve the formative assessment.	3
Students did take notes on scientific methods, do the students have discussions within the groups?	5
Science	2
After having students read from the text, consider having pairs work together for a few minutes on the activities in the text. You could circulate while they work in order to monitor their progress.	7
MS. __identifies critical information and students respond accordingly.	3
Students should not be listening to music during class time. Electronic devices are only to be used for academic purposes!	2
Students should not have hoods on in class!	2
Lots of enthusiasm and encouragement, great job!	4
How do you determine that all students understand the content that is being taught?	5

Data Collection

The quantitative data representing interpersonal behavior traits, IBS responses were collected from participating Osceola County School administrators during IBS administration sessions in May of 2017. Administrator IBS responses were used to develop a full IBS profile for each administrator. While IBS responses were identifiable to the researcher during the onset of the analysis, no responses with information that would identify the respondent will be reported to any school district officials or the public. The researcher employed a numbered covered sheet for IBS responses on which the respondent wrote their name and other demographic information. Once IBS scores were paired to feedback scores, the researcher removed the cover sheet and used only the number system. The cover sheets were secured and retained in the event that it became necessary to repeat the process of pairing feedback scores to IBS scores. Although the researcher developed an individual IBS profile for each administrator, the raw IBS scale scores for each administrator, on each scale of the IBS, were used during the quantitative analysis. Feedback data were gathered from the district; for the purposes of this study, feedback examined was limited to that which was entered into the district's evaluation system, My Professional Growth System which uses the Marzano framework.

Approval to conduct this study was sought from and granted by the University of Central Florida's Institutional Review Board (IRB). Approval to conduct this study in Osceola County was granted by The Superintendent of The School District of Osceola County, Dr. Debra Pace, and the district's Research, Evaluation, and Accountability

department. Documents attesting to approval from both bodies can be found in the

appendixes of this study.

Data Analysis

This study was guided by an interest in determining what relationship exists

between the various personality traits held by school administrators, and the degree to

which they hold them, and the quality of the feedback provided to the teachers assessed

by that administrator.

Quantitative data of the eight IBS scale scores for assertiveness, the seven IBS

scores for aggressiveness, the three IBS scale scores for relationships, and the three IBS

scale scores under validity, feedback rubric scores, and respondent demographic

information (years in administration, age, and gender) was collected, analyzed, and

reported for the descriptive statistical measures of range, mean, and standard deviation.

Multiple regression analysis and Pearson r were calculated, using The Statistical Package

for Social Sciences (SPSS). Multiple regression analysis was performed to determine the

relationship between the seven scale scores on the aggressiveness scales of the IBS and

the quality of the feedback given by the administrator. Multiple regression analysis was

conducted a second time to determine the relationship between the eight scale scores on

the assertiveness scales of the IBS and feedback. Multiple regression analysis was

conducted a third time to determine the relationship between the three scale scores on the

relationship scales of the IBS and the quality of feedback provided. For the purpose of the

multiple regression, an average of all feedback scores were used in the analysis. The

findings of the multiple regression were used to answer the research questions. Pearson correlation analysis was performed along with each of the four multiple regressions to determine the relationship between individual behavior trait measures within each group (i.e. conflict avoidance within the relationship grouping) and feedback quality. While the research questions are formulated such that the multiple regression analysis truly answers what is being asked, the Pearson *r* provides additional insight into the relationship between the various behavior traits measured by the IBS and feedback quality.

According to Steinberg (2011) Pearson correlation analysis provides "a measure of the linear relationship between two variables that have both been measured on at least an interval level" (p. 432). Formula for calculating Pearson *r* is:

$$rxy = \frac{N\sum XY - \sum X \sum Y}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}}$$

In the above formula X equals the subject's raw score on variable x, Y equals the subject's raw score on variable y, and N equals sample size (Steinberg, 2011).

A multiple regression is the act of "predicting a criterion outcome based on more than one predictor variable simultaneously" (Steinberg, 2011, p. 488). For this study the IBS scale scores will serve as the predictor, or independent, variables while the feedback score will serve as the dependent variable. By determining the extent to which the IBS scale scores can serve to predict the feedback score, the researcher will gain an understanding of the relationship between the IBS scores and the feedback score, this will allow the researcher to address the IBS score to feedback score relationship component of the research questions.

82

Regression is "the process of predicting the dependent variable from the independent variable" (Steinberg, 2011, p. 487). The dependent variable in this study is the mean score for all of the feedback given by an administrator during the 2015-16 school year. If this study was using only one predictor variable, simple linear regression or linear prediction could have been conducted; however, because this study is using multiple predictor variables, multiple regression is necessary. According to Steinberg (2011) "Predicting a criterion outcome based on more than one predictor variable simultaneously is called multiple regression (p. 488). The researcher chose to use more than one predictor variable because "rarely is an outcome adequately predicted by a single predictor" (Steinberg, 2011, p. 489). The predictor variables in this study are the scale scores for the 21 interpersonal behavior traits measured by the Interpersonal Behavior Survey (IBS). The equation for multiple regression is:

$$Y'=b_1X_1+b_2X_2+b_3X_3+\ldots b_nX_n=a$$

In the equation n=the number of predictor variables, in this study that will range from as few as three for research questions 3 and 4 and as many as eight for research question 1. Y' or Y prime is predicted criterion score, each X is a predictor variable and each b is the weight for that variable; "the weight tells the relative importance of that variable in predicting the criterion's variance" (Steinberg, 2011, p. 489).

Given that study seeks to understand what, if any, relationship exists between a set of independent variables—interpersonal behavior trait scores—and a dependent variable—feedback quality as represented by mean feedback score—multiple regression

analysis was deemed to be the best statistical method of arriving at an understanding of how interpersonal behavior traits relate to feedback quality.

Based upon the results of the analysis discussed above, the researcher determined it would be beneficial to conduct power analysis to provide additional insight into the findings of this study. Specific discussion of why this analysis was deemed useful in the context of this study is provided in chapter 4; however, given the treatment of the other types of analysis and calculation performed in this study in this particular chapter, it was deemed appropriate to provide some explanation of power analysis here.

In this study, power analysis was performed after data were collected and after other analysis had been performed; however, "power analysis is normally conducted before the data collection. The main purpose underlying power analysis is to help the researcher to determine the smallest sample size that is suitable to detect the effect of a given test at the desired level of significance" ("Statistical Power Analysis", 2018). Given that the researcher had no ability to control the sample size in this study, conducting power analysis at the outset was not deemed useful. In light of the findings, it was conduct in an attempt to avoid making a type II—retaining the null hypothesis when it ought to be rejected. G*Power software, from the University of Dusseldorf, was used to conduct the power analysis calculations for this study.

Summary

In this chapter, the purpose of this study and research questions were restated. The selection of subjects, instrumentation, data collection procedures, and data analysis processes were also discussed. The use of additional analysis, power analysis, to better understand the findings of this study was also discussed. Results of the data analysis will be presented in chapter 4.

Table 7. Research Questions, Research Methods, Variables, Data Sources, and Data Analysis

Research Questions	Research Method	Variables	Data Sources	Analysis
1. What, if any, relationship exists between the assertiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?	Quantitative	Independent: The 8 IBS assertiveness scale scores. Dependent: feedback scores.	IBS response scores and feedback scores.	Multiple regression
2. What, if any, relationship exists between the aggressiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?	Quantitative	Independent: The 7 IBS aggressiveness scale scores. Dependent: feedback scores.	IBS response scores and feedback scores	Multiple regression
3. What, if any, relationship exists between the relationship scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?	Quantitative	Independent: The 3 IBS relationship scale scores. Dependent: feedback scores.	IBS response scores and feedback scores	Multiple regression

4. What is the relationship between scores above 70T on the Denial, Infrequency, and Impression Management scales—all scores that render the aggressiveness, assertiveness, and relationship portions of the respondent's survey unreliable—and the feedback provided by the administrators with those scores?	Quantitative	Independent: IBS denial, infrequency, and impression management scale scores. Dependent: feedback scores.	IBS response scores and feedback scores.	Multiple regression

CHAPTER 4
PRESENTATION AND ANALYSIS OF DATA

Introduction

The purpose of this study was to determine if—and to what extent—interpersonal behavior traits, and the degree to which an administrator possesses them, can explain the well-documented phenomenon of inflated teacher evaluation scores. In order to accomplish this goal, the researcher administered the Interpersonal Behavior Survey (IBS) to all school-based administrators in Osceola County who consented to participating in the study. The survey results rendered interpersonal behavior trait scale scores for each participating administrator. At the same time the researcher also read feedback comments given by the participating school-based administrators and scored them using a feedback-scoring rubric (Rafalski, 2015). Using IBM's SPSS software, the researcher then conducted multiple regression analysis to determine what, if any, relationship exists between interpersonal behavior traits and feedback.

This chapter will present the results of this analysis. First, descriptive statistics for the study participants—including demographic and IBS score data—will be provided. This will be followed by a discussion of the results of each multiple regression analysis performed to answer each of the four research questions. Following the discussion of the analysis of findings related to each of the four research questions, a discussion of the important additional findings of this study will be provided. The discussion will include both the SPSS statistical output as well as a narrative discussion of what the statistical output means.

<u>Overall Descriptive Statistics</u>

Of the fifty-four school-based administrators—out of approximately 120 in the district— who agreed to participate in this study, four declined to include any identifying information with their Interpersonal Behavior Survey (IBS) answer sheet; due to the inability to connect these IBS score profiles to any feedback, these four respondents are included in table 7 below but are not included in any of the calculations conducted for this study. Three other administrators did not enter any teacher observations for the 2015-16 school year. Because level 1 on the feedback rubric represents no feedback, these three administrators were given an average feedback score of 1. Nine of the administrators—four men and five women—had a score above the reliability cut-off on one or two of the validity scores. While IBS score information for these nine administrators is represented on table 7 below, they are only included in the analysis to answer question 4—the unreliable nature of their IBS profiles was the reason for their exclusion. This reduces the population for questions one through three to 41— representing 32% of the total school-based administrator population.

The validity scales of the IBS—denial, infrequency, and impression management—serve as a means of determining the validity of the full interpersonal behavior profile developed by the IBS. As Mauger & Adkinson (1980) note, "This group of scales reflects test-taking attitudes….The person's attitude toward the inventory affects scores on the assertive, aggressive, and relationship scales. The validity scales indicate whether a profile is interpretable" (p. 2). The denial scale measures the hesitancy to admit to common but socially undesirable foibles: Things such as using profanity and making

89

fun of others. The infrequency scale indicates "the tendency to endorse items that less than 10% of the normative sample endorsed" (Mauger & Adkinson, 1980, p. 2). The impression management scale provides an indication of the respondent attempting to provide answers that reflect well on him or her as a person, rather than the honest answer. A more detail discussion of impression management can be found under the definition of terms in chapter one of this study. Because the denial. Infrequency, and impression management scales reflect an individual's willingness to be dishonest in responding to IBS items, a high scale on any one of these three scales call into questions all the responses given by an individual respondent. Their scores on other items cannot be trusted, because they demonstrated a tendency to be, repeatedly, dishonest in responding to IBS items.

When looking at the IBS scale scores, there are a few important things about the IBS scores to keep in mind. First, the scores on the validity scores—denial, infrequency, and impression management—can render the profile unreliable. Scores above 6 on the denial scale, above 6 for men and 5 for women on the infrequency scale, and 21 on the impression management scale render the profile unreliable. As mentioned above, nine study participants had scores outside the acceptable range on one or two of these scales—none had scores outside the acceptable range on the infrequency scale. Additionally, while higher scores on the assertiveness scales—general assertiveness, self-confidence, initiating assertiveness, defending assertiveness, frankness, praise, requesting help, and refusing demands—are considered a positive, high scores on the aggressiveness scales—general aggressiveness, hostile stance, expression of anger, disregard for rights, verbal

aggressiveness, physical aggressiveness, and passive aggressiveness—are considered a negative attribute. Table 7 below provides the highest possible IBS scale score for each scored trait, along with the mean study participant score, and the accompanying standard deviation. Because the highest possible IBS scale score varies by gender on a few of the scales, this information is presented here broken-down by gender.

Table 8. Participant IBS Scores by Gender

IBS Scale	Highest Possible Score on the IBS Instrument		Mean Participant Score		Standard Deviation	
	Male	Female	Male	Female	Male	Female
Denial	9	9	4.09	3.65	1.73	1.69
Infrequency	10	8	.43	.45	.93	.68
Impression Management	29	29	17.9	17.13	4.15	3.58
General Aggressiveness	38	36	6.43	5	5.74	3.99
Hostile Stance	23	23	5	3.83	4.2	3.01
Expression of Anger	19	19	1.95	2.28	2.38	2.15
Disregard for Rights	11	10	1.52	1.35	1.69	1.49
Verbal Aggressiveness	14	14	2.43	2.43	1.93	1.55
Physical Aggressiveness	13	12	1.86	1.41	1.71	1.23
Passive Aggressiveness	32	31	5.19	5.3	4.53	3.92
General Assertiveness	55	55	36.76	40.45	8.33	5.95
Self-Confidence	16	16	9.23	10.2	3.54	3.27
Initiating Assertiveness	17	17	10.81	11.68	2.6	2.14
Defending Assertiveness	18	18	12.67	14.18	3.95	2.5
Frankness	12	12	6.71	7.45	2.67	2.39
Praise	9	9	5.1	6.2	2.19	1.86
Requesting Help	7	7	4.14	4.3	2.1	2.03
Refusing Demands	6	6	4.1	4.58	1.48	1.01
Conflict Avoidance	22	22	11.48	10.83	5	4.09
Dependency	23	23	8.9	8.1	4.55	4.3
Shyness	23	23	8.3	8.1	5.74	6.81

It should be noted here that of the 50 school-based administrators who consented to participating in the study and provided identifying information that allowed the researcher to pair their IBS scores with feedback scores— 4 study participants gave no identifying information—nine had scores above the cut-off point on one or more of the validity scales—seven on either the denial or impression management scale and two on both the denial and impression management scale. Because these scores render the respondents other IBS scale scores unreliable, the IBS and feedback scores for these nine individuals were only used in the analysis toward answer research question four, where including them was necessary, their scores were otherwise excluded from the analysis due to the unreliable nature of their full IBS profile. This means that 17% of the 54 respondents had to be excluded due to their denial or impression management score, or a combination of them.

Study participants ranged in age from 33 to 66—33 to 61 in male participants and 34 to 66 in female participants. Their years of experience in an administrative position— assistant principal and principal—ranged from 2 years to 32 years—3 years to 32 years in male participants and 2 years to 30 years in female participants. Table 9 below provides an overview of study participant age and years of experience in administration.

Table 9. Study Participant Age & Experience.

	Age		Years in Administration	
	Mean	Standard Deviation	Mean	Standard Deviation
Male	48.19	8.55	12.4	7.5
Female	45.03	15.18	10.4	7.17
All	43.95	16.27	10.54	7.49

Of the 5,586 feedback comments provided to teachers by the 54 school-based administrators who participated in this study, the largest portion, 3,201 comments, comprising 57.3% of comments rated, were a recount of classroom events, a level 3 on the rating rubric. With 1,087 comments, 19.5% of the total, the next largest portion of the feedback fell into the category of no feedback—meaning the comment box was left blank or that the administrator conducted no teacher observations at all—that is scored as a level 1 on the rubric. Only 192 comments—or 3.4% of the total—were rated as specific targeted feedback, a level 7on the rating rubric. It is important to note here that level 6 on the rubric—standardized feedback—was not scored for any of the comments in this study as it refers to comments that can be copied and pasted from a bank of available comments in a platform that was no longer in use in the district during the time period from which comments for this study were collected. A complete breakdown of the scoring for all the feedback examined in this study is provided in table 10 below.

Table 10. Breakdown of Study Feedback Scores.

Feedback Rubric Score	Count	Percent of Total
1 No feedback	1087	19.5
2 Unrelated feedback	269	4.8
3 Recount of events	3201	57.3
4 General affirmation	483	8.6
5 Reflective question	354	6.3
6 Standardized	0	0
7 Specific Targeted	192	3.4

Research Question 1

What, if any, relationship exists between the assertiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

In order to arrive at an answer to this question a multiple linear regression analysis was performed using the mean feedback score for each of the school-based administrator as the dependent variable and the eight scores for each of the interpersonal behaviors that fall under the assertiveness section of the IBS as the predictors. These eight behaviors are: general assertiveness, self-confidence, initiating assertiveness, defending assertiveness, frankness, praise (giving/receiving), requesting help, and refusing demands. This analysis included 41 of the 54 study participants. Four were excluded due to an inability to pair their IBS scale scores with feedback scores—due to incomplete IBS answer sheets. Another nine were excluded due to scores outside the acceptable range on one or more of the IBS validity scales. Further explanation of the validity scales and why they disqualify respondents from the analysis, can be found at the beginning of this chapter. Table 11 below provides the mean feedback score and standard

deviation for each of the eight assertiveness scale scores. Reference table 8 for a context

for these scores.

Table 11. Mean & standard deviation for assertiveness & feedback scores

	Mean	Standard deviation
Feedback	2.81	.89
General Assertiveness	37.41	7.00
Self-Confidence	9.41	3.27
Initiating Assertiveness	11.00	2.43
Defending Assertiveness	12.82	3.18
Frankness	6.97	2.48
Praise	5.51	2.09
Requesting Help	4.12	1.98
Refusing Demands	4.29	1.32

Prior to conducting the multiple regression analysis to understand the relationship

between feedback quality and the eight interpersonal behavior traits measured on the

assertiveness scales of the IBS as a group, the researcher conducted Pearson correlation

analysis to observe the relationship between each assertiveness scale score, individually,

and feedback quality.

Table 12. Pearson Correlations for assertiveness scores & feedback quality

IBS Assertiveness Scale	r	p
General Assertiveness	.045	.391
Self-Confidence	-.190	.117
Initiating Assertiveness	.100	.268
Defending Assertiveness	.133	.203
Frankness	.167	.148
Praise	-.088	.292
Requesting Help	-.159	.161
Refusing Demands	-.024	.440

The Pearson correlation analysis shows weak correlations between five of the assertiveness scales and feedback quality. The five are: self-confidence (r=-.190), initiating assertiveness (r=.100), defending assertiveness (r=.133), frankness (r=.167), and requesting help (-.159). The other three, Praise (-.088), Refusing Demands (r= -.024), and General Assertiveness (r= .045) are weak. Two of the five with weak correlations—self-confidence and requesting help—are negatively correlated with feedback quality. As mentioned above, this is noteworthy as the opposite relationship would be expected. None of the assertiveness scale score correlations with feedback quality are significant, all are at p > .05.

A multiple linear regression was calculated to predict feedback quality based on eight assertiveness interpersonal behavior trait scores. Results of the multiple linear regression indicate there is no collective significant effect between the eight IBS assertiveness scale scores and feedback quality, $(F_{(8,32)}= .815, p = .595)$, with an R^2 of .169. The multiple correlation between the set of predictor variables, the IBS assertiveness scale scores, and the criterion, average feedback, was found to be R= .411. According to Steinberg (2011) "A correlation of .82 is strong, while a correlation of .13 is weak" (p. 422). Given that the correlation here, .411, falls near the middle of this range from weak to strong, the correlation here can be considered moderate. The weighted combination of the predictor variables explains approximately 17% of the variance of the average feedback score (R^2=.169). The assertiveness scale scores as a predictive model have a significance of p=.595. With this p value, it could be said that the researcher has failed to predict significance, and that the null hypothesis—that random happenstance is

just as likely to account for the variance in feedback quality as the levels of assertiveness are—should be accepted here; however, given the high-moderate correlation, the possibility of a type II error is present here. A type II error occurs when a null hypothesis is accepted when it ought not have been. The small sample size for the analysis, 41 participants, is a leading reason to be concerned about a type II error here; a more detailed explanation of why this is the case will be provided below. Table 13 below provides an overview of the analysis discussed above.

Table 13. Results of multiple regression using assertiveness scores & feedback quality.

N	R	R^2	df	F	p
41	.411	.169	8	.815	.595

In an attempt to further understand R and p values observed here, the researcher made us of Gpower software to conduct power analysis. Based on this analysis, a study population of 82—double that of this study—would have had significant findings.

Coefficients are weights; represented by *B* for the regression equation. They tell how much the dependent variable is expected to increase—if the weight is positive—or decrease—if the weight is negative—when that independent variable increases by one, holding all other independent variables constant. Table 13 below shows that of the eight assertiveness scale scores, refusing demands (-.138) and self confidence (-.220) are the only coefficients of note. It should be noted here that both of these IBS assertiveness traits show a negative relationship, as they increase, feedback quality decreases. This is noteworthy as it is the opposite of what would be expected, given that the assertiveness

98

traits are expected to improve one's ability to successfully engage in interpersonal communication. As was true of the assertiveness scores as a predictive model, none of the coefficients are significant; with all at p >.05. Table 14 below provides a complete accounting of the coefficients for each of the assertiveness scales scores.

Table 14. Coefficients for assertiveness scale scores and feedback quality

IBS Assertiveness Scale	B	t	p
Constant	2.11	2.27	.03
General Assertiveness	.076	.840	.407
Self-Confidence	-.220	-1.433	.162
Initiating Assertiveness	-.032	-.345	.733
Defending Assertiveness	.006	.067	.947
Frankness	.047	.546	.589
Praise	.030	.249	.805
Requesting Help	.070	.506	.616
Refusing Demands	-.138	-.898	.376

While the correlation between the IBS assertiveness scale scores as predictor variables and feedback quality as dependent variable is moderate at R=.411, the accompanying value of p=.595 leads the researcher to accept the null hypothesis. As noted in the discussion of the power analysis above, the researcher acknowledges there is a possibility of type II error here and that the failure to predict significance observed here could be due to the size of the study population and not due to a lack of true significance in the assertiveness scores of the individual providing feedback. The results of the multiple regression show that the assertiveness scales scores were not significant predictors of feedback quality.

Research Question 2

What, if any, relationship exists between the aggressiveness scores, as measured by the

Interpersonal Behavior Survey (IBS), of school administrators and the feedback they

provide to the teachers they supervise?

In order to arrive at an answer to this question a multiple linear regression

analysis was performed using the mean feedback score for each of the school-based

administrator as the dependent variable and the seven scores for each of the interpersonal

behaviors that fall under the aggressiveness section of the IBS as the predictors. These

seven behaviors are: general aggressiveness, hostile stance, expression of anger,

disregard for rights, verbal aggression, physical aggression, and passive aggression. This

analysis included 41 of the 54 study participants. Four were excluded due to an inability

to pair their IBS scale scores with feedback scores—due to incomplete IBS answer

sheets. Another nine were excluded due to scores outside the acceptable range on one or

more of the IBS validity scales. Further explanation of the validity scales and why they

disqualify respondents from the analysis, can be found at the beginning of this chapter.

Table 15 below provides the mean feedback score and standard deviation for each of the

seven aggressiveness scale scores. Reference table 8 for a context for these scores.

Table 15. Mean & Standard deviation for aggressiveness & feedback scores

	Mean	Standard deviation
Feedback	2.81	.89
General Aggressiveness	6.12	5.09
Hostile Stance	4.8	3.8
Expression of Anger	2.63	2.37
Disregard for Rights	1.53	1.73
Verbal Aggression	2.7	1.8
Physical Aggression	1.56	1.39
Passive Aggressiveness	5.87	4.32

Prior to conducting the multiple regression analysis to understand the relationship between feedback quality and the seven interpersonal behavior traits measured on the aggressiveness scales of the IBS as a group, the researcher conducted Pearson correlation analysis to observe the relationship between each aggressiveness scale score, individually, and feedback quality.

Table 16. Pearson Correlations for aggressiveness scores & feedback quality

IBS Assertiveness Scale	r	p
General Aggressiveness	.16	.15
Hostile Stance	.15	.17
Expression of Anger	.03	.42
Disregard for Rights	.09	.27
Verbal Aggression	.13	.2
Physical Aggression	.17	.14
Passive Aggression	-.08	.13

The Pearson correlation analysis shows weak correlations between four of the aggressiveness scales and feedback quality. The four are: general aggressiveness (r= .16), hostile stance (r= .15), verbal aggression (r= .13), and physical aggression (r= .17). All of these four are positively correlated with feedback quality; expression of anger (r= .03)

101

and disregard for rights (r= .09) are also positively correlated, albeit very weakly. This is noteworthy as the opposite relationship would be expected—given that high quantities of aggressiveness traits are not desirable. Interestingly, only passive aggressiveness shows the negative correlation (r = -.08) that is expected with aggressiveness traits. None of the aggressiveness scale score correlations with feedback quality are significant, all are at p > .05.

A multiple linear regression was calculated to predict feedback quality based on eight aggressiveness interpersonal behavior trait scores. Results of the multiple linear regression indicate there is no collective significant effect between the seven IBS aggressiveness scale scores and feedback quality ($F(7,33)= .508$, $p = .821$), with an R^2 of .097. The multiple correlation between the set of predictor variables, the IBS aggressiveness scale scores, and the criterion, average feedback, was found to be R=.312. According to Steinberg (2011) "A correlation of .82 is strong, while a correlation of .13 is weak" (p. 422). Given that the correlation here, .312, falls near the middle of this range from weak to strong, the correlation here can be considered moderate-to-weak. The weighted combination of the predictor variables explains approximately 0.9% of the variance of the average feedback score($R^2 = .097$). The aggressiveness scale scores as a predictive model have a significance of p =.821. With this p value, it could be said that the researcher has failed to predict significance, and that the null hypothesis—that random happenstance is just as likely to account for the variance in feedback quality as the levels of aggressiveness are—should be accepted here; however, given the weak-to-moderate correlation, the possibility of a type II error is present here. A type II error

occurs when a null hypothesis is accepted when it ought not have been. The small

sample size for the analysis, 41 participants, is a leading reason to be concerned about a

type II error here; a more detailed explanation of why this is the case will be provided

below. Table 17 provides an overview of the analysis discussed above.

Table 17. Results of multiple regression using aggressiveness scores & feedback quality

N	R	R^2	df	F	p
41	.312	.097	7	.508	.821

In an attempt to further understand R and p values observed here, the researcher

made us of Gpower software to conduct power analysis. Based on this analysis, a study

population of 93—11 more than double that of this study—would have had significant

findings. This finding regarding the size of the study population is why the possibility of

a type II error exists—as discussed above.

Coefficients are weights; represented by *B* for the regression equation. They tell

how much the dependent variable is expected to increase—if the weight is positive—or

decrease—if the weight is negative—when that independent variable increases by one,

holding all other independent variables constant. As was true of the aggressiveness scores

as a predictive model, none of the coefficients are significant; with all at p >.05. Table 18

below provides a complete accounting of the coefficients for each of the aggressiveness

scales scores.

Table 18. Coefficients for aggressiveness scale scores and feedback quality

IBS Assertiveness Scale	B	t	p
Constant	2.7	8.18	.00
General Aggressiveness	.02	.2	.84
Hostile Stance	.00	.02	.98
Expression of Anger	-.05	-.61	.54
Disregard for Rights	-.05	-.32	.74
Verbal Aggression	.13	.75	.45
Physical Aggression	.09	.56	.57
Passive Aggression	-.06	-1.37	.17

While the correlation between the IBS aggressiveness scale scores as predictor variables and feedback quality as dependent variable is weak-to-moderate at R=.312, the accompanying value of p=.821 leads the researcher to accept the null hypothesis. As noted in the discussion of the power analysis above, the researcher acknowledges there is a possibility of type II error here and that the failure to predict significance observed here could be due to the size of the study population and not due to a lack of true significance in the aggressiveness scores of the individual providing feedback. The results of the multiple regression show that the aggressiveness scales scores were not significant predictors of feedback quality.

<u>Research Question 3</u>

What, if any, relationship exists between the relationship scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

In order to arrive at an answer to this question a multiple linear regression analysis was performed using the mean feedback score for each of the school-based

administrator as the dependent variable and the three scores for each of the interpersonal

behaviors that fall under the relationship section of the IBS as the predictors. These three

behaviors are: shyness, conflict avoidance, and dependency. This analysis included 41 of

the 54 study participants. Four were excluded due to an inability to pair their IBS scale

scores with feedback scores—due to would-be participants leaving demographic sections

of their IBS answer sheets blank. Another nine were excluded due to scores outside the

acceptable range on one or more of the IBS validity scales. Further explanation of the

validity scales and why they disqualify respondents from the analysis, can be found at the

beginning of this chapter. Table 19 below provides the mean feedback score and standard

deviation for each of the three relationship scale scores. Reference table 8 for a context

for these scores.

Table 19. Mean & Standard deviation for relationship & feedback scores

	Mean	Standard deviation
Feedback	2.81	.89
Conflict Avoidance	11.36	4.06
Dependency	8.92	4.5
Shyness	9.39	6.13

Prior to conducting the multiple regression analysis to understand the relationship

between feedback quality and the three interpersonal behavior traits measured on the

relationship scales of the IBS as a group, the researcher conducted Pearson correlation

analysis to observe the relationship between each relationship scale score, individually,

and feedback quality.

Table 20. Pearson Correlations for relationship scores & feedback quality

IBS Assertiveness Scale	r	p
Conflict Avoidance	-.24	.06
Dependency	.07	.32
Shyness	.15	.16

The Pearson correlation analysis shows weak correlations between two of the three relationship scales and feedback quality. The two are: conflict avoidance (r = -.24) and shyness (r = .15). Dependency (r = .07) is positively correlated; however, the correlation is very weak. While shyness is positively correlated—feedback quality increases as shyness increases—conflict avoidance shows both a stronger correlation and, as would be expected with this trait, the correlation is negative—as conflict avoidance increases, feedback quality decreases. While it is true that none of the relationship scale score correlations with feedback quality are significant, all are at p > .05, it should be noted that conflict avoidance (p = .06) is very close to significance. Of the 18 traits discussed thus far in the study thus far, none come closer than conflict avoidance to have a significant relationship with feedback quality.

A multiple linear regression was calculated to predict feedback quality based on three interpersonal relationship trait scores. Results of the multiple linear regression indicate there is no collective significant effect between the three IBS relationship scale scores and feedback quality (F(3,37)= 2.07, p = .12), with an R^2 of .38. The multiple correlation between the set of predictor variables, the IBS relationship scale scores, and the criterion, average feedback, was found to be R=.38. According to Steinberg (2011) "A correlation of .82 is strong, while a correlation of .13 is weak" (p. 422). Given that the correlation here, .38, falls near the middle of this range from weak to strong, the

correlation here can be considered moderate. The weighted combination of the predictor variables explains approximately 14% of the variance of the average feedback score ($R^2 =$ .144). The relationship scale scores as a predictive model have a significance of $p = .12$. With this p value, it could be said that the researcher has failed to predict significance, and that the null hypothesis—that random happenstance is just as likely to account for the variance in feedback quality as the levels of relationship skills are—should be accepted here; however, given the moderate correlation, the possibility of a type II error is present here. A type II error occurs when a null hypothesis is accepted when it ought not have been. The small sample size for the analysis, 41 participants, is a leading reason to be concerned about a type II error here; a more detailed explanation of why this is the case will be provided below. Table 21 provides an overview of the analysis discussed above.

Table 21. Results of multiple regression using relationship scores & feedback quality.

N	R	R^2	df	F	p
41	.38	.14	3	2.07	.12

In an attempt to further understand R and p values observed here, the researcher made us of Gpower software to conduct power analysis. Based on this analysis, a study population of 69—28 more than that of this study—would have had significant findings. This finding regarding the size of the study population is why the possibility of a type II error exists—as mentioned above.

Coefficients are weights; represented by *B* for the regression equation. They tell how much the dependent variable is expected to increase—if the weight is positive—or

decrease—if the weight is negative—when that independent variable increases by one, holding all other independent variables constant. Table 22 below shows that the coefficient weight for each of the three relationship scale traits are rather small with conflict avoidance (-.08) being the largest. While both shyness and dependency are not significant here, $p > .05$ for both, conflict avoidance is significant at $p = .03$.

Table 22. Coefficients for relationship scale scores and feedback quality

IBS Assertiveness Scale	B	t	p
Constant	2.93	6.01	.00
Conflict Avoidance	-.08	-2.15	.03
Dependency	.05	1.64	.1
Shyness	.03	1.35	.18

While the correlation between the IBS relationship scale scores as predictor variables and feedback quality as dependent variable is moderate at $R = .38$, the accompanying value of $p = .12$ leads the researcher to accept the null hypothesis. As noted in the discussion of the power analysis above, the researcher acknowledges there is a possibility of type II error here and that the failure to predict significance observed here could be due to the size of the study population and not due to a lack of true significance in the relationship scores of the individual providing the feedback. The results of the multiple regression show that the relationship scales scores were not significant predictors of feedback quality.

Research Question 4

What is the relationship between scores above 70T on the Denial, Infrequency, and Impression Management scales—all scores that render the aggressiveness, assertiveness, and relationship portions of the respondent's survey unreliable—and the feedback provided by the administrators with those scores?

In order to arrive at an answer to this question a multiple linear regression analysis was performed using the mean feedback score for each of the school-based administrator as the dependent variable and the three scores for each of the interpersonal behaviors that fall under the validity section of the IBS as the predictors. These three behaviors are: denial, infrequency, and impression management. Because this question seeks to explore the relationship between scores outside the range of reliability—normed across all three scales as 70T—and feedback quality, only 9 of the 54 study participants were used for this analysis. 41 participants were excluded due to not having a score outside the range of reliability on any of the three validity scales. Another 4 would-be participants were excluded due to incomplete demographic information on their IBS answer sheets that rendered it impossible to match their IBS profiles to feedback scores. Table 23 below provides the mean feedback score and standard deviation for each of the three validity scale scores. Reference table 8 for a context for these scores.

Table 23. Mean & Standard deviation for relationship & feedback scores

	Mean	Standard deviation
Feedback	2.76	.86
Denial	5.88	1.26
Infrequency	.11	.33
Impression Management	22.66	2.12

Prior to conducting the multiple regression analysis to understand the relationship between feedback quality and the three traits measured on the validity scales of the IBS as a group, the researcher conducted Pearson correlation analysis to observe the relationship between each validity scale score, individually, and feedback quality.

Table 24. Pearson Correlations for validity scores & feedback quality

IBS Assertiveness Scale	r	p
Denial	-.16	.33
Infrequency	-.18	.31
Impression Management	.6	.04

The Pearson correlation analysis shows weak correlations between two of the three validity scales and feedback quality. The two are: denial ($r = -.16$) and infrequency ($r = -.18$). Both are negatively correlated; as the trait increases, the quality of feedback declines. Interestingly, impression management shows a much larger correlation ($r = .6$) and is positively correlated. As the impression management score increases, the quality of feedback increases. Additionally, while the correlations for both denial and infrequency are not significant, both are at p .05, the correlation between impression management and feedback quality is significant at $p = .04$.

A multiple linear regression was calculated to predict feedback quality based on the three validity scale scores of the IBS. Results of the multiple linear regression indicate there is no collective significant effect between the three IBS validity scale scores, when the score is above the accepted level for validity, and feedback quality, $(F(3,5)= 1.19, p = .4)$, with an R^2 of .41. The multiple correlation between the set of predictor variables, the IBS relationship scale scores, and the criterion, average feedback, was found to be R = .54. According to Steinberg (2011) "A correlation of .82 is strong, while a correlation of .13 is weak" (p. 422). Given that the correlation here, .64, falls above the middle of that range, closer to strong, the correlation can be said to be moderate-to-strong. The weighted combination of the predictor variables explains approximately 41% of the variance of the average feedback score ($R^2 = .41$). The validity scale scores as a predictive model have a significance of p = .4. With this p value, it could be said that the researcher has failed to predict significance, and that the null hypothesis—that random happenstance is just as likely to account for the variance in feedback quality as the validity scale scores are—should be accepted here; however, given the moderate-to-strong correlation, the possibility of a type II error is present here. A type II error occurs when a null hypothesis is accepted when it ought not have been. The small sample size for the analysis, 9 participants, is a leading reason to be concerned about a type II error here; a more detailed explanation of why this is the case will be provided below. Table 25 provides an overview of the analysis discussed above.

Table 25. Results of multiple regression using validity scores & feedback quality.

N	R	R^2	df	F	p
9	.64	.41	3	1.19	.4

In an attempt to further understand R and p values observed here, the researcher made us of Gpower software to conduct power analysis. Based on this analysis, a study population of 20—11 more than that of this study—would have had significant findings. It should be noted here that this would require finding 11 more participants with a score above the cut-off for reliability on one or more of the validity scales. Simply adding 11 more participants to the study would not guarantee this outcome—this is evidenced by the number of participants (44) that did not exceed the reliability threshold on any of the three of the validity scales. There was a tenth participant who exceeded the reliability threshold on two of the three validity scales, but this would-be participant did not include sufficient demographic information to connect the IBS profile to feedback scores.

Coefficients are weights; represented by *B* for the regression equation. They tell how much the dependent variable is expected to increase—if the weight is positive—or decrease—if the weight is negative—when that independent variable increases by one, holding all other independent variables constant. Table 26 below shows that the coefficient weight for the validity scales range from small (Denial, .06) to rather large (infrequency, -.58). None of the validity traits show significance, all are p > .05.

Table 26. Coefficients for validity scale scores and feedback quality

IBS Assertiveness Scale	B	t	p
Constant	-3.57	-.84	.43
Denial	.06	.25	.8
Infrequency	-.58	-.65	.53
Impression Management	.26	1.75	.13

While the correlation between the IBS validity scale scores as predictor variables and feedback quality as dependent variable is strong at R = .64, the accompanying value of p = .4 leads the researcher to accept the null hypothesis. As noted in the discussion of the power analysis above, the researcher acknowledges there is a possibility of type II error here and that the failure to predict significance observed here could be due to the size of the study population and not due to a lack of true significance in the validity scores of the individual providing the feedback. The results of the multiple regression show that the validity scale scores above 70T were not significant predictors of feedback quality.

<u>Additional Analysis Findings</u>

Having conducted analysis to respond to each of the research questions that guide this study, the researcher opted to conduct additional analysis using two of the pieces of demographic data collected on the IBS answer sheet: age and years as a school administrator. This analysis included 48 of the 54 study participants. Four were excluded due to incomplete IBS answer sheets that made it impossible to pair the IBS profile with feedback scores. Two others were excluded because they did not enter their age on the

IBS answer sheet. Table 27 below provides the mean feedback score and standard

deviation for age and years of experience as a school administrator.

Table 27. Mean & Standard deviation for age, years of admin experience & feedback

	Mean	Standard deviation
Feedback	2.83	.86
Years of Administrative experience	11.04	7.42
Age	47.95	7.85

Prior to conducting the multiple regression analysis to understand the relationship

between feedback quality and years of administrative experience and age as a

combination, the researcher conducted Pearson correlation analysis to observe the

relationship between years of administrative experience and age, individually, and

feedback quality.

Table 28. Pearson Correlations for demographic factors & feedback

Demographic Factor	r	p
Years in Administration	-.31	.01
Age	-.31	.01

The Pearson correlation analysis shows moderate correlations between both of the

demographic factors. Both are negatively correlated; as the demographic characteristic

increases, the quality of feedback declines. The correlation between both demographic

factors and feedback quality is significant with both at $p < .05$.

A multiple linear regression was calculated to predict feedback quality based on

years of administrative experience and age. Results of the multiple linear regression

indicated that there was a collective significant effect between the age, years of

administrative experience, and feedback quality, ($F(2,45)= 3.02$, $p = .05$), with an R^2 of .11. The multiple correlation between the set of predictor variables, years of administrative experience and age, and the criterion, average feedback, was found to be $R=.34$. According to Steinberg (2011) "A correlation of .82 is strong, while a correlation of .13 is weak" (p. 422). Given that the correlation here, .34, falls in the middle of that range, it can be said to be a moderate correlation. The weighted combination of the predictor variables explains approximately 11% of the variance of the average feedback score ($R^2=.11$). The demographic characteristics as a predictive model have a significance of $p=.05$. Here the researcher will reject the null hypothesis and accept that a significant relationship does exist between years of administrative experience, age, and feedback quality.

Table 29. Results of multiple regression using demographics & feedback quality

N	R	R^2	df	F	p
48	.34	.11	2	3.02	.05

Coefficients are weights; represented by *B* for the regression equation. They tell how much the dependent variable is expected to increase—if the weight is positive—or decrease—if the weight is negative—when that independent variable increases by one, holding all other independent variables constant. Table 29 below shows that the coefficient weight for the demographic traits—years in administration and age.

Table 30. Coefficients for years in administration & age with feedback

Demographic Trait	B	t	p
Constant	4.03	4.69	.00
Years in Administration	-.02	-1.03	.3
Infrequency	-.02	-.95	.34

The correlation between years of administrative experience and age as predictor variables and feedback quality as dependent variable is moderate at R=.34, and with the accompanying value of p=.05 the researcher rejects the null hypothesis. The results of the multiple regression show that years of experience in administration and age are significant predictors of feedback quality.

<div align="center">Summary</div>

In this chapter, data were analyzed in order to reach answers to four research questions that focused on the relationship between interpersonal communication traits and the quality of the feedback provided to teachers by their supervising school-based administrator.

Research question 1 asked what the nature was, if any, of the relationship between interpersonal traits listed under assertiveness and feedback quality. Assertiveness traits are considered an asset, so higher levels of these traits are, generally, positive. The traits listed under assertiveness on the IBS are general assertiveness, self-confidence, initiating assertiveness, defending assertiveness, frankness, praise (giving/receiving), requesting help, and refusing demands. A multiple linear regression calculation was performed, the findings were such that no significant relationship could be said to exist between the

assertiveness scales as a group, nor between any individual assertiveness scale, and the quality of feedback.

Research question 2 asked what the nature of the relationship between interpersonal traits listed under aggressiveness on the IBS and feedback quality. Aggressiveness traits are considered an impediment to productive interactions, so higher levels of these traits are, generally, negative. The traits listed under aggressiveness on the IBS are general aggression, hostile stance, expression of anger, disregard for rights, verbal aggression, physical aggression, and passive aggression. A multiple linear regression calculation was performed, the findings were such that no significant relationship could be said to exist between the aggressiveness scales as a group, nor between any individual aggressiveness scale, and the quality of feedback provided by the administrator.

Research question 3 asked if any relationship exists between the three traits rated on the relationship scale of the IBS and the quality of the feedback provided by school-based administrators who posses the traits. The three relationship scales, conflict avoidance, dependency, and shyness, generally carry negative connotations and would be expected—at higher levels—to be a hindrance to productive interpersonal relations. A multiple linear regression calculation was performed, the findings were such that no significant relationship could be said to exist between the relationship scales as a group, nor between any individual assertiveness scale, and the quality of feedback; although, conflict avoidance came closer than any of the other traits measured by the IBS to being significant, ultimately, here it could not be proven to be of statistical significance.

117

Research question 4 sought to determine if high scores on any of the validity scales of the IBS had a relationship with feedback quality. The validity scales of the IBS are used to determine whether or not the full IBS profile is reliable. Scores above a set cut-off (70T) on each of the three scales, denial, infrequency, and impression management, render the rest of the profile unreliable. A multiple linear regression was performed to answer this question. The results of the multiple regression were the same as was found in questions 1, 2, and 3: No significant relationship could be said to exist between validity scale scores above the cut-off and feedback quality.

An additional multiple linear regression calculation was performed, this time using two pieces of demographic information, age and years of administrative experience. This calculation proved more fruitful than the others: A significant relationship was found. Both years of experience in administration and age were found to have significant, negative relationships with feedback quality. The older and more tenured the administrator, the lower the feedback quality.

Table 31. Research Questions, Variables, Data Sources, Methods of Analysis & Results

Question #	Research Question	Variables	Data Sources	Method of Analysis	Results
1	What, if any, relationship exists between the assertiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?	Independent Scores on the 8 assertiveness scales of the IBS. Dependent Mean feedback score.	Responses to the Interpersonal Behavior Survey (IBS). Feedback given during 2015-16 school year and scored by the researcher using a feedback scoring rubric.	Multiple linear regression.	No significant findings at p>.05.
2	What, if any, relationship exists between the aggressiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?	Independent Scores on the 7 aggressiveness scales of the IBS. Dependent Mean feedback score.	Responses to the Interpersonal Behavior Survey (IBS). Feedback given during 2015-16 school year and scored by the researcher using a feedback scoring rubric.	Multiple linear regression.	No significant findings at p>.05.

Table 31. Research Questions, Variables, Data Sources, Methods of Analysis & Results

Question #	Research Question	Variables	Data Sources	Method of Analysis	Results
3	What, if any, relationship exists between the relationship scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?	<u>Independent</u> Scores on the 3 relationship scales of the IBS. <u>Dependent</u> Mean feedback score.	Responses to the Interpersonal Behavior Survey (IBS). Feedback given during 2015-16 school year and scored by the researcher using a feedback scoring rubric.	Multiple linear regression.	No significant findings at p>.05.
4	What is the relationship between scores above 70 on the validity scales and the feedback provided by the administrators with those scores?	<u>Independent</u> Scores on the 3 validity scales of the IBS. <u>Dependent</u> Mean feedback score.	Responses to the Interpersonal Behavior Survey (IBS). Feedback given during 2015-16 school year and scored by the researcher using a feedback scoring rubric.	Multiple linear regression.	No significant findings at p>.05.

CHAPTER 5
SUMMARY, DISCUSSION, AND RECOMMENDATIONS

Introduction

This chapter is intended to further the conversation begun in the chapters that precede it. This chapter provides a summary of the overall study, further discussion of the research findings—both those directly related to the research questions that guided this study and the additional findings that were made independent of the research questions. This chapter also provides an overall discussion of the study findings, a discussion of what implications the findings of this study have for practitioners, and recommendations for future research. The summary of the study includes an abbreviated versions of the statement of the problem, purpose of the study, theoretical framework, research questions, and research methodology. Following the summary of the study, the next section of this chapter will focus on a discussion of the findings of the study; this section will begin with a review and discussion of the findings related to the four research questions and then delve into a discussion of the additional findings. This discussion of the findings will be followed by a discussion of the overall findings of this study. This will be followed be a section exploring what, if any, implications this study has for current practice. The exploration of the implications for practice will be followed by recommendations for future research. The recommendations for future research will be followed be a conclusion to the study.

Summary of the Study

Since the beginning of the modern accountability in education movement began—approximately 1983 with the publication of A Nation at Risk and reaching its legislative-pinnacle in 2001 when No Child Left Behind (NCLB) was enacted—there has been an emphasis on student proficiency and, more importantly, on the effectiveness of teachers to move their students toward reaching proficiency, especially in reading and mathematics.

Studies (Mela, 2013, Pace, 2015, Rafalski, 2015, and Butler, 2017) have shown that, when compared to the achievement of their students, the evaluation scores, or effectiveness ratings, of teachers are inaccurate, and they show signs of being inflated. As Mela (2013) states: "This means that [Brevard County] is still experiencing to a lesser degree; however, what Weisberg et al. (2009) called the Widget Effect where all teachers' classroom effectiveness was judged the same with no distinction being made between the good and the poor teachers" (p. 137). In order for a poor teacher's score to match that of a good teacher, there must be score inflation occurring. In short, school administrators who evaluate teacher effectiveness, principals and assistant principals, are inflating teacher effectiveness scores, this inflation becomes apparent when teacher effectiveness scores are compared to student achievement scores.

Despite the existance of several studies that establish the inflated nature of teacher effectiveness ratings (Mela, 2013, Pace, 2015, Rafalski, 2015, and Butler, 2017), prior to this study, no study had sought to explore interpersonal behavior traits as a means of accounting for the observed inflation of teacher effectiveness ratings. By employing the Interpersonal Behavior Survey (IBS), this study explored the relationship between the interpersonal

behavior traits held by administrators and the feedback they provided to the teachers they supervised. This was done in attempt to arrive at an understanding of or explanation for the observed inflation of teacher effectiveness ratings. Marx's conflict theory provided the theoretical foundation for this study. Marx's theory states that those with power and resources—the bourgeois—are in a constant state of conflict with those who lack power and resources—the proletariat. In the context of the schoolhouse and inflated effectiveness ratings, Marx's conflict theory can be understood in the following way: By inflating scores, administrators encourage docility in those they supervise. By allowing the proletariat to think they have won the battle, the bourgeois win the war. It costs the administrator—the bourgeois—nothing in terms of real resources or power to tell a mediocre teacher that he or she is a good teacher. Indeed, this act is likely to bring about docility in the worker and allow the bourgeois to maintain power without any real conflict. Inflating effectiveness ratings preclude conflict.

Four research questions were developed to explore this theory and to gain an understanding of the relationship between interpersonal behavior traits and feedback:

1. What, if any, relationship exists between the assertiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

2. What, if any, relationship exists between the aggressiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

3. What, if any, relationship exists between the relationship scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

4. What is the relationship between scores above 70 on the Denial, Infrequency, and Impression Management scales—all scores that render the aggressiveness, assertiveness, and relationship portions of the respondent's survey unreliable—and the feedback provided by the administrators with those scores?

In order to arrive at answers to these questions, the researcher administered the IBS to all Osceola school-based administrators—principals and assistant principals—who consented to participating in this study. This resulted in an IBS profile consisting of 21 scores—a score for each trait measured by the IBS. At the same time, the researcher read and scored all the feedback given by the school-based administrators, who consented to participate in this study, during the 2015-16 school year. The scoring was done using a feedback scoring rubric (Rafalski, 2015). The researcher then conducted multiple regression analysis to determine what, if any, relationship exists between interpersonal behavior traits as measured by the IBS and the quality of the feedback given to teachers by their administrators.

<u>Discussion of the Findings</u>

Research Question 1

What, if any, relationship exists between the assertiveness scores, as measured by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

Upon completing multiple regression analysis, the researcher noted a weak-to-moderate correlation existed between the eight IBS assertiveness scores combined and the average feedback score of the administrators in the study. The eight assertiveness scales are general assertiveness, self confidence, initiating assertiveness, defending assertiveness, frankness, praise (accepting and giving), requesting help, and refusing demands. While the correlation is close to what is commonly accepted as a moderate correlation, the model was not significant; however, power analysis showed that with a population size of 82, double the population available to the researcher in this study, significance could be found.

Research Question 2

What, if any, relationship exists between the aggressiveness scores, as measured   by the Interpersonal Behavior Survey (IBS), of school administrators and the feedback they provide to the teachers they supervise?

Upon completing multiple regression analysis, the researcher noted a correlation between the seven aggressiveness scales and the quality of feedback provided that fell between weak and moderate. As was the case with the assertiveness scales, this correlation

was not significant. The seven aggressiveness scales of the IBS are general aggressiveness, hostile stance, expression of anger, disregard for rights, verbal aggression, physical aggression, and passive aggressiveness. As was true with assertiveness in question one, power analysis revealed that a larger study population—93 participants this time—would have resulted in significant findings. Again, the size of the study population precludes definitively ruling out the aggressiveness traits measured by the IBS as influencing the quality of feedback given. It should be noted, however, that the correlation between the assertiveness traits—those deemed to be positive attributes—was greater than that between the quality of feedback and aggressiveness traits—the traits deemed less desirable.

Research Question 3

What, if any, relationship exists between the relationship scores, as measured by   the Interpersonal Behavior Survey (IBS), of school administrators and the   feedback they provide to the teachers they supervise?

The three IBS relationship scales are conflict avoidance, dependency, and shyness. As was true with the assertiveness scales and the aggressiveness scales, the multiple regression analysis resulted in a correlation between the relationship scales and the quality of feedback that fell between weak and moderate. No significant relationship was found. Power analysis was again conducted, this time a population of 69, 28 more participants than were available to the researcher in this study, was determined to be where significance would have been found. Additionally, of note with the relationship scales is conflict avoidance. Of all the traits measured by the IBS, conflict avoidance showed a stronger correlation to feedback quality

126

than any other single trait. Additionally, as would be expected the correlation found was negative, as conflict avoidance increased, feedback quality decreased. While no significant relationship was found, the correlation between conflict avoidance and feedback quality was closer to .05—it was .06—than any other relationship observed in this study.

Research Question 4

What is the relationship between scores above 70 on the Denial, Infrequency, and Impression Management scales—all scores that render the aggressiveness, assertiveness, and relationship portions of the respondent's survey unreliable—and the feedback provided by the administrators with those scores?

The study population, 50 total participants, yielded 9 respondents with a score above 70 on one or more of the three validity scales. The three validity scales are denial, infrequency, and impression management. In order to arrive at an answer for this research question, the researcher conducted multiple regression analysis using the validity scale scores of the 9 respondents who exceeded 70 on one or more of the scales and the average feedback scores for these 9 participants. The results were a correlation that comes close to strong, and, again, a significance well above .05. The researcher once again conducted power analysis. This analysis suggests a population of 20 would allow for significant findings. While this is an increase of 11 more participants, simply increasing the study population by 11 would not be sufficient here as there is no way of guaranteeing the 11 new participants would have a score above 70 on one or more of the validity scales. Indeed, given what these scales measure—a tendency to deny possessing common yet undesirable traits, a tendency to less

127

frequently identify with such traits, and a tendency to shape ones answers in order to cultivate a better impression of one's self—it would be expected that the occurrence of scores above 70 on any one or more of these scales would be rare and undesirable among the ranks of leadership as these traits are not ones typically associated with quality leadership; this will be discussed further later in this chapter.

Additional Findings

While it cannot be said that this study definitively linked behavioral traits to the quality of feedback and thus found one possible explanation for the observed inflation of teacher effectiveness ratings, this study did arrive at three findings that are noteworthy. The first is the overall quality of feedback which confirms the findings of previous studies. The second is the number of respondents excluded from the study due to scores above the cut-off on one or more of the three validity scales, and the third is that as years of experience in school administration and age increase, the quality of feedback provided decreases.

Of the 59,967 pieces of feedback scored for her study Rafalski (2015) found that 13,122 pieces of feedback were scored as a one, or no feedback given; 138 pieces of feedback were scored as unrelated feedback, a two on the scale; 34,682 pieces of feedback were found to be a recount of observed events, a three on the scale; 3,581 pieces of feedback were rated as a four on the scale, general affirmation; 2,778 pieces of feedback were found to be reflective questions, a five on the scale; 1,442 pieces of feedback were found to be standardized feedback—using boilerplate feedback copied and pasted from the observation platform, a six on the scale; and 4,224 pieces of feedback were found to be specific and targeted feedback, a seven on the scale (Rafalski, 2015, p. 107). The reason the breakdown of

the feedback scores in Rafalski's study are worth noting here is, despite being separated by three years and occurring in two different school districts, the breakdown of the feedback in Rafalski (2015) is very similar to the breakdown found in this study, especially when looking at the percentage of feedback accounted for by each rubric level. In this study 1087 pieces of feedback were rated at a level one; 269 pieces were rated at a level two; 3201 were rated a three; 483 were rated a four; 354 were rated a five; none were rated a six as the specific observation platform was not used; and 192 pieces were rated a seven. The table below provides a side-by-side comparison of the ratings.

Table 32. Feedback score breakdown Rafalski, 2015 & Walker, 2018

Level	Feedback type	Rafalski, 2015	Walker, 2018
		% of Total Feedback	
1	No Feedback	21.8	19.5
2	Unrelated Feedback	0.2	4.8
3	Recount of observed events	57.8	57.3
4	General Affirmation	5.9	8.6
5	Reflective Question	4.6	6.3
6	Standardized Feedback	2.4	N/A
7	Specific targeted feedback	7	3.4

While all of these percentages are similar, the most notable similarity is that in both studies the rating of 3, a recount of observed events, accounts for roughly 57% of the scored feedback. This is especially noteworthy as it is the most benign in that it is neither positive nor negative. It is the most innocuous feedback an administrator can give while still giving some form of feedback that does relate to the observed lesson; however, it offers nothing in the way of actual useful feedback for the teacher. Given that the goal of feedback

is to improve instruction, the fact the majority of the feedback simply relates what was observed is indeed troubling.

The study found that as the age and years of administrative experience a study participant had increased, the quality of the feedback they offered decreased. This finding contradicts the conventional wisdom of experience with a task or practice with a task leading to improved performance. This finding parallels those of Ling (2014) who found that as individuals progressed along the school leadership continuum from graduate students to assistant principals to principals their moral decision making declined.

Ling (2014) asserts that "…additional analyses of this study's data further revealed serious trouble in the principalship. Current principals' failure to reach the higher levels of Postconventional reasoning and decision-making bodes extraordinarily poorly for those students and stakeholders looking toward principals to lead the way" (p. 134).

Finally, of the 54 would-be participants in this study, 10 had a score above the acceptable cut-off on one or more of the three validity scales of the IBS. One of these individuals was also excluded from the study due to failing to provide enough demographic information to pair the IBS profile with feedback scores. 18% of the would-be study population was excluded due to a high score on the denial or impression management scale. According to Mauger and Adkinson (1980) the denial scale "indicates a hesitancy to admit to common but socially undesirable weaknesses and feelings. This includes such behaviors as making fun of others, swearing, and procrastinating" (p. 2). The impression management scale relates to a desire to "create a favorable impression" of oneself (Mauger & Adkinson, 1980, p. 2). Essentially, high scores on either of these scales expose a degree of intentional

dishonesty on the part of the respondent that renders all of their responses to the items on the IBS unreliable. John Brooks Slaughter asserts that "Individuals become leaders only when their behaviors, thoughts, and values, as seen or perceived by those they would lead dictate that they are deserving of their roles" (Brooks Slaughter, 2012, p.85). The degree of willful dishonesty noted here, while compatible with the findings of other studies (Strenth, 2013 & Ling, 2014) hardly meets the benchmark for leadership established by Brooks Slaughter.

<div align="center">Overall Discussion of the Findings</div>

In an attempt to arrive at an understanding of what these findings mean, let us begin by removing The Lake Wobegon Effect as a possible explanation for the observed teacher effectiveness rating inflation. While the Garrison Keillor informed theory may appear to be the perfect explanation, none of the other facts observed substantiate this theory. If, as this theory posits, everyone is above average, there would be a glut of feedback scored as praise, not to mention the accompanying above-average student performance. Furthermore, The Lake Wobegon Effect simply provides a name for what is being observed, it does not provide a root-cause explanation for the observed phenomenon.

Given what the literature (Hattie & Timperly, 2007, Hattie, 2009, Sadle, 1989, Shute, 2008, and DeNisi & Kluger, 2000) says about the importance feedback to improving performance, it is no stretch to say that providing feedback to teachers is an essential job for school leaders. Indeed, the ability to provide instruction-improving feedback is increasingly an expectation of school leaders (Jerald, 2012). Despite this fact, this study and others (Rafalski, 2015) have shown that the majority of the feedback being provided to teachers by

their administrators is lacking any actionable information. Why, if feedback is so essential to improvement, is the quality of it so poor and continually coupled with inflated teacher effectiveness ratings? Deception. Under the umbrella of interpersonal communication theory lies interpersonal deception theory. The theory posits that individuals engage in deception for a number of reasons, one reason is personal gain (Buller & Burgoon, 1996). In the theoretical framework of this study, the idea was put forth that the inflated effectiveness ratings are a means of control exercised by administrators to pacify and appease teachers. The inflated effectiveness ratings are, given their failure to align with the reality of student performance, a deception engaged in by administrators for their gain. What is gained by this deception? Well, as one administrator who declined to be named in this study stated, "There's a teacher shortage, so if you can't keep teachers, eventually your boss will wonder why he should keep you." The simple answer here is job security. To quote Michael Douglas as President Andrew Shepard in the film The American President, "I was so busy keeping my job, I forgot to do my job." The abdication of duty being asserted here is not without precedent in the research, studies (Vitton & Wasonga, 2009, Strenth, 2013, and Ling, 2014) have, troublingly, shown lacking moral decision making to not be an aberration among school leaders. Denying teachers the feedback necessary to improve their craft can hardly be seen as anything other than a failing on the part of their administrator. The strong negative correlation observed between conflict avoidance and feedback quality observed in this study—a relationship that very nearly proved significant—furthers the idea that the what is being observed is a deception, one that allows administrators to maintain power while avoiding conflict.

What is being advocated for here is not for administrators to begin berating, chastising, or castigating teachers. Given the current teacher shortage, such an approach would be foolish to say the least. However, if feedback is not going to be of useful quality and effectiveness ratings are going to in inflated, well, what is the point of having either? Additionally, as noted behavioral biologist and former collaborator of B. F. Skinner and Konrad Lorenz, Karen Pryor (1999) notes there is also risk in continually issuing false praise: "False or meaningless praise, however, is soon resented, even by tiny children, and loses any power to reinforce" (p. 15). So, administrators ought not get caught in the deceit of empty feedback and inflated, false, effectiveness ratings and begin engaging in the difficult work of earnest feedback and accurate effectiveness ratings; even when this means delivering less than positive commentary to teachers. Again, Pryor (1999) says it best, "reprimands," she notes, "are a necessary part of existence" (p. 19).

Implications for Practice

The implications this study has for current practice are fairly direct: The quality of the feedback being given to teachers by their administrators has to improve. While the study was unable to definitively link interpersonal behavior traits to feedback quality, it did confirm— as have other studies (Rafalski, 2015)—that the overall quality of the feedback being given to teachers by their administrators is poor, lacking in specificity and anything actionable. It is impossible to say whether or not addressing the problem of poor feedback quality will have any impact of the primary concern of this study, the observed inflation of teacher effectiveness ratings, it is fair to speculate that improving feedback will improve instruction.

Perhaps by improving instruction through improved feedback, the problem of inflated teacher effectiveness ratings will be addressed by bringing actual teacher effectiveness up to close the gap between what student performance tells us about teacher effectiveness and how administrators rate teacher effectiveness. While it would be valuable to have a thorough explanation for the inflation phenomenon, ultimately, bringing student performance up to match the current effectiveness ratings would be the most beneficial outcome, and improving the quality of feedback provided to teachers by their administrators can certainly serve as an important piece in that process.

Recommendations for Future Research

Power analysis continually suggested that a larger study population would have allowed the researcher in this study to achieve statistically significant findings, as such, the first suggestion for future research would be to replicate this study with a large population. With that end in mind, the complete data set—all the IBS scale scores and average feedback scores—for all 50 study participants is included as an appendix to this study.

Additionally, of the 21 items measured on the IBS, the one that came closest to significance was conflict avoidance. This behavior fits with the theoretical underpinnings of this study, Marxian conflict theory, it also fits with the type of feedback overwhelmingly observed in both this study and others (Rafalski, 2015), the very benign recounting of observed classroom events. This feedback that includes exclusively the facts of what occurred while providing no qualitative commentary, neither finding fault or praise, is less likely to lead to conflict than more effective feedback that specifically notes a deficit and

134

provides routes for addressing it. While not entirely true, the old adage, you cannot argue with the facts, seems appropriate here. Given all of this, a study focusing specifically on conflict avoidance in school-based administrators would certainly be useful in extending this conversation.

Finally, given the prevalence of mindset, determination and drive in contemporary thinking around success (Dweck, 2006, Pink, 2011, and Duckworth, 2016), it is being advised here that this work be continued on looking at risk taking and risk avoiding behaviors. While this study has framed the discussion using conflict avoidance—a term used by the IBS—it is easy to understand risk taking and risk avoiding as closely related to conflict avoidance. Those willing to take risks often face conflict in the form of resistance from the play-it-safe folk. Additionally, Dweck (2011) frames an unwillingness to take risks as evidence of a fixed mindset, a trait not desirable for those seeking sustainable success and lasting impact.

Conclusion

This study sought to explore the observed phenomenon of inflated teacher effectiveness ratings by determining what, if any, relationship exists between interpersonal behavior traits and the quality of administrator-to-teacher feedback. Although this study did not prove that statistically significant relationships exist between any of the measured interpersonal behavior traits and administrator-to-teacher feedback, it is not without merit. The study did show a significant relationship between administrator age, years of experience in administration, and feedback quality—specifically that feedback quality declined as age

135

and years of experience in administration increase. This study also confirmed the findings of prior studies (Rafalski, 2015) showing that the overall quality of administrator-to-teacher feedback is poor. Finally, when considered as part of a larger conversation including prior studies (Mela, 2013, Strenth, 2013. Ling, 2014, Pace, 2015, Rafalski, 2015, and Butler, 2017) this study—or rather these studies—suggest that the inflation of teacher effectiveness ratings is not caused by any particular observation system or protocol, but by the people doing the rating.

APPENDIX A OSCEOLA COUNTY RESEARCH APPROVAL

THE SCHOOL DISTRICT OF OSCEOLA COUNTY, FLORIDA

817 Bill Beck Boulevard • Kissimmee• Florida 34744-4492
Phone: 407-870-4600 • Fax: 407-870-4010 • www.osceolaschools.net

SCHOOL BOARD MEMBERS

District 1 – Jay Wheeler
407-973-4141
District 2 – Kelvin Soto – Chair
407-361-2462
District 3 – Tim Weisheyer
407-361-0235
District 4 – Clarence Thacker
407-870-4009
District 5 – Ricky Booth – Vice Chair
407-870-4009

Superintendent of Schools
Dr. Debra P. Pace

March 29, 2017

Robert Walker
723 Hayden Lane
Orlando, FL 32804

Dear Mr. Walker:

This letter is to inform you that we have received your request to conduct research in our School District. Based on the description of the research you intend to conduct, I am pleased to inform you that you may proceed with your work as you have outlined. Please be advised that this approval is based on the understanding that a school's participation is completely voluntary and left to the discretion of each building administrator. Please also be advised that the district office will not be able to assist you with any aspect of your research (e.g. sending emails, obtaining data, locating students, providing addresses, etc.).

Finally, be reminded that all information obtained for the purpose of your research must be dealt with in the strictest of confidentiality. At no time is it acceptable to release any student or staff identifiable information. Upon completion of your research, please provide our office with a copy of your results.

I wish you the best of luck in your future endeavors. If I can be further assistance, please do not hesitate to contact me.

Sincerely,

Janine Jarvis, Director
Research, Evaluation & Accountability

Student Achievement – Our Number One Priority
Districtwide Accreditation by the Southern Association of Colleges and Schools
An Equal Opportunity Agency

APPENDIX B UCF INSITITUIONAL REVIEW BOARD RESEARCH APPROVAL

University of Central Florida Institutional Review Board Office of Research & Commercialization 12201 Research Parkway, Suite 501 Orlando, Florida 32826-3246 Telephone: 407-823-2901 or 407-882-2276 www.research.ucf.edu/compliance/irb.html

Approval of Human Research

UCF Institutional Review Board #1 FWA00000351, IRB00001138

To: Date: Dear Researcher: On 05/01/2017 the IRB approved the following human participant research until 04/30/2018 inclusive:

From:

Robert C Walker May 01, 2017

Type of Review: Project Title:

Investigator: IRB Number: Funding Agency: Grant Title:

Research ID:

UCF Initial Review Submission Form Expedited Review Interpersonal Behavior Traits and Their Relationship to

Administrator-to-Teacher feedback: A Quantitative Study. Robert C Walker SBE-17-13101 N/A

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form cannot be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu .

If continuing review approval is not granted before the expiration date of 04/30/2018, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate. Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

140

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual. On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

_____ _____ _____

_____ _____ _____

Renea Carver

Signature applied by Renea C Carver on 05/01/2017 03:49:05 PM EDT IRB Coordinator

APPENDIX C STUDY INFORMED CONSENT

Interpersonal Behavior Traits and Their Relationship to Administrator-to-Teacher Feedback: A Quantitative Study
Informed Consent for an Adult in a Non-Exempt Research Study

Principal Investigator: Robert C. Walker, MFA

Faculty Advisor: Barbara Murray, PhD

Investigational Site(s):   The School District of Osceola County

Introduction: Researchers at the University of Central Florida (UCF) study many topics. To do this we need the help of people who agree to take part in a research study. You are being invited to take part in a research study about interpersonal behavior traits and administrator-to-teacher feedback which will include about 130 Osceola County school administrators. You have been asked to take part in this research study because you are a school-based administrator in Osceola County. You must be 18 years of age or older to be included in the research study.

The person doing this research is Robert Walker a doctoral student in UCF's educational leadership program in the school of teaching, learning, and leadership in the college of education and human performance. Because the researcher is a doctoral student he is being

guided by Dr. Barbara Murray, a UCF faculty advisor in the college of education and human performance.

What you should know about a research study:

- Someone will explain this research study to you.
- A research study is something you volunteer for.
- Whether or not you take part is up to you.
- You should take part in this study only because you want to.
- You can choose not to take part in the research study.
- You can agree to take part now and later change your mind.
- Whatever you decide it will not be held against you.
- Feel free to ask all the questions you want before you decide.

Purpose of the research study: The purpose of this study is to determine what, if any,

relationship exists between interpersonal behaviorial traits and administrator-to-teacher

feedback.

What you will be asked to do in the study: You will be asked to complete the 272 item

Interpersonal Behavior Survey (IBS).

Location: Per an arrangement with The School District of Osceola County (SDOC), IBS responses will be collected during currently schedule School District of Osceola County administrator meetings in May 2017. The meeting location will be set by SDOC.

Time required: We expect that you will be in this research study for the 45 minutes

necessary to complete the IBS.

Risks: As all IBS responses will be de-identified at the onset of and for the duration of the study, the risk to respondendents is less than minimal; however, if you feel uncomfortable you may discountinue participation at any time. Additionally, no identifiable information will be used in reporting the findings. While the researcher may use terms such as "a middle school" or "a high school", no specific school names or other potentially identifying information will be used.

Benefits: While respondents are likely find the insight gained from reflecting on their IBS scores to be benefical to their professional practice, especially in their interactions with colleagues, there are no expected financial benefits to you for taking part in this study.

Compensation or payment: There is no compensation or other payment to you for taking part in this study.

Anonymous research: This study is anonymous. That means that no one, not even members of the research team, will know that the information you gave came from you. You will be identified only by respondent number, 2015-16 worksite, and 2015-16 department(s) or grade level(s) supervised. No names or employee identification numbers will be used in this study.

Study contact for questions about the study or to report a problem: If you have questions, concerns, or complaints, or think the research has hurt you, talk to Robert Walker, Doctoral Student, Education Leadership Program, College of Education and Human Performance, (407) 617-8439 or Dr. Barbara Murray, Faculty Supervisor, Department of Education Leadership at 407-823-1473 or by email at Barbara.Murray@ucf.edu.

IRB contact about your rights in the study or to report a complaint: Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (UCF IRB). This research has been reviewed and approved by the IRB. For information about the rights of people who take part in research, please contact: Institutional Review Board, University of Central Florida, Office of Research & Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246 or by telephone at (407) 823-2901. You may also talk to them for any of the following:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.

- You want to get information or provide input about this research.

Withdrawing from the study:

If you decide to leave the study, contact the investigator so that the investigator can remove your responses from data to be analyzed. You will need your respondent number in order for the research to identify and remove your responses. The sponsor can also end the research study early. We will tell you about any new information that may affect your health, welfare or choice to stay in the research.

APPENDIX D FEEDBACK SCORING RUBRIC

No Feedback (Level one)	Unrelated Feedback or General Statement Provided (Level 2)	Recount of Classroom Events (Level 3)- Justification for rating.	General Affirmation Statement (Level 4)	Reflective Question (Level 5)	Standardized Feedback Provided (Level 6)	Specific Targeted Feedback Provided (Level 7)
	2-The message is Unintelligible	3-Recap has several different components (sometimes statement of percent of students being monitored or desired effect). Sometimes actually gives examples of what is wrong with no suggestion. You did this or that, teacher did this or that, students did this or that, I observed this or that...	4-General praise. Good job, great job, excellent job, I liked, I loved, WOW!	5- Asks the teacher a question.	6- Examples: How might you adapt and create new strategies for chunking content into digestible bites that address unique student needs and situations? How might you expand your monitoring to involve more students? What are you learning about your students as you adapt and create new strategies? In addition to monitoring students by the use of choral responses, how else can you monitor students when chunking information?	7-Language like: 1. Reference to Resource Library or Reflective Teacher 2. Maybe try.... Or You might want to try.... 3. Consider.... 4. Recommendation.... Or I would recommend.... 5. Suggestion.... Or I suggest.... 6. It might be a good idea... 7. You should.... 8. This would have been good or great if... 9. To move to a higher level, do this_____. 10. Think about..... 11. I want you to.... 12. Next time.... 13. Always.... 14. Be sure to... or Make sure you..... 15. Doing this_____would have been more effective. 16. You need to.... 17. Continue to... 18. Remember this_____. 19. Do this_____. 20. This_____is a good strategy. 21. I would like to see.... 22. Coaching idea...

APPENDIX E EXAMPLE FEEDBACK TEXT

Text included here was randomly selected from all the feedback scored for this study. The text here is intended to serve as an example of the feedback scored for this study. It is not intended to be representative of all the feedback scored for this study.

1. The students were in charge of tracking their own progress.

2. You are a rock star!

3. Development measures/ factors was a concept your students struggled with based on an FRQ. This lesson allowed your students to dig and gain a deeper understanding. I noticed on the groups larger than two there was not equal engagement among all the students. Some students were not engaged at all. I wonder how could you increase the engagement and participation for ALL of your students?

4. The students were not only reflecting on their own learning but served as a resource to one another.

5. Students read aloud and discussed the text.

6. Students did take notes on scientific methods, do the students have discussions within the groups?

7. During her whole group lesson, Ms. ____ provided opportunities for students to discuss concepts in partners. During partner discussions, students were engaged and focused. I would recommend providing some additional opportunities for them to talk. During the whole group reading, five students were off task and not paying attention. Providing these opportunities would provide some additional accountability and purpose.

8. Teacher began lesson with a quick launch activity focusing on producers and consumers. Students worked in partners regarding pictures and connections being made, teacher spoke with each group of students to monitor their understanding, she asked clarifying questions.   Your classroom is very well organized and your students are very aware of your expectations, your room runs very smoothly. Your students were using science academic vocabulary throughout the entire lesson. It was great when you reminded them to use the new science vocabulary words that they were learning about today. Students then worked in small groups on developing their food chain.

9. Working with 2 students at teacher table., 3 students on the computer. Working on reading with emotion using exclamation points. Monitored the two students at the table.   Did not notice teacher watching the 3 the computer.

10. Bell work is a review of last three days. Students stood up to their answer choice when checking them as a class. Students working on chapter 13 scavenger hunt review as groups. Teacher circulating among groups.

11. Mr. ____ was introducing procedures for microscope use and safety in preparation for looking at cells. He had a student demonstrate handling procedures and how to carry the microscope. He then asked a couple of other students various questions on procedures for using a microscope.   In your monitoring of content- you asked a couple of students for answers. How could you monitor differently to ensure all of your students are understanding the content?

12. This strategy was clearly called for as the students were watching the movie 'The Hobbit.' The teacher remained at his desk and the movie ran continuously for the entire length of my observation in the room. There were no goals/objectives visible on the board to indicate the purpose for viewing. Some students were watching the move while a few had out cell phones and others were talking or drawing. It did not seem, from their actions, that a purpose for viewing had been articulated. When asked, one student indicated that they were not given any specific tasks to engage in while viewing the movie, nor any indication of what tasks/activities would follow. The students recently read the novel, and it is my assumption that there will be some activity to follow that will tie the two together. However, this strategy would have allowed the teacher to support students in making connections between the two versions of the story.

13. Secret listen 2d and 3D How can I tell you the secret. You need to listen.   Here is the secret Get ready Don't forget the secret information.

14. I could not tell what was going on in the classroom. There was too much down time with students gluing homework into the agenda. Some were on the computer, others walking around the room. Five minutes until iii but most kids were not doing much that entire time. Some started reading when they had nothing to do. Let's talk about engaging students during the math block and differentiating instruction with teacher group being part of those minutes.

15. You effectively questioned the students to elaborate on their responses.

16. Reading aloud of the story Macbeth. Mr.___ stopped and explained what was happening with the witches. Moved on to an audio recording to continue the story. No essential question or objectives on the board. Need to have them to guide the lesson of the day. 2 students had their heads down and were not engaged in the lesson.

17. Students were given an index card with a number 1-10 and had to order themselves. Students sitting clapped if in the correct order. It seemed 2 of them got it, but they'll get it with more practice!

18. Great lesson and excellent rapport with your students. You have established a very trusting and secure academic environment. Excellent work.

19. Students were arranged in groups of 3 and 4 to discuss the content of the lesson. The teacher asked questions about theme. Students seemed attentive. Suggestion: Perhaps have students work collaboratively on the question and individual respond. Although many students participated it would have been difficult to be sure of the students' mastery.

20. Review of story structure was the learning objective, specifically point of view and characters. ELA standard .2.RL 2.5 was being reviewed. Review a bit scattered due to set up and inadvertent technical issues during actual review. Use of interactive technology a plus and the interjection of comprehension questions were limited to more compliant engagement vs. authentic. Just some basic procedural things in place would have made for a more effective lesson.

21. Teacher begins class by asking students where they left off in reading Twilight. What is the purpose for reading the next pages or chapters? Be very overt to the students on why you are doing the activities that you are doing. Three students responded out loud as to where they left off reading, as the rest of the students waited to begin reading. While one student read aloud, teacher sat in a chair at the front, and students were not asked to do anything but follow along. Three students had their head down, while two others were looking about the classroom. What was the expectation of reading aloud? Earlier, the students explained that they were working on vocabulary, but there were no words listed on board, or given to the students for them to be looking for or paying attention to. Is the reading of Twilight purely for ACT vocabulary practice? Are students to be listening for comprehension/understanding of the story? Character development? Teacher mentioned that good comprehension is as if there is a movie playing in your head and you can see what is happening in pictures...are students able to write and describe or draw what they see as happening or important to the story or characters? How is what they are reading being represented? What ACT/FSA skills could students be practicing while reading the text?

22. I would like to see something here that really gets students excited and understanding what it is that you are about to teach them and why you are teaching it.

23. Use of white erase board.

24. I thought this was algebra, your working on multiplying fractions. We are 12 weeks into the school year. This is algebra 1. You need to move on.

25. Reading - I can determine the character traits in Judy Moody that led to the sequence of the story. 8:55 - Review what are character traits, apply motivation to Judy Moody story 8:58 Students round robin read and teacher asks questions 906 student confused on motivation...what was her motivation? 9:09 review scale & ask students on levels 1 & 2 if can move sticker ; review sequence of story 920 return to seats to share with partner assigned question 9:34 focus on Stink point of view 9:36 continue to work on questions 9:39 independently - sequence next Judy Moody story 9:49 sequence of new story - thinking map 9:58 Do you see why sequence of story is important? 10:00 homework Sequence by telling the beginning and end of story & then fill in the events?? How is a Venus Flytrap a pet? When did the students fill in the reflection sheet?

LIST OF REFERENCES

Altman, I. & Taylor, D. (1973). *Social penetration: The development of interpersonal relationships.* New York: Holt, Rinehart, and Winston.

Ashley, S. A. (2015). *The effects of the marzano observation system training on the self-efficacy of teacher observers* (Doctoral dissertation). Retrieved from http://etd.fcla.edu.ezproxy.net.ucf.edu/CF/CFE0005758/Dissertation_SamuelAlvinAshleyIII_finalsubmission.pdf

Axley, S. R. (1984). Managerial and organizational communication in terms of the conduit metaphor. *The Academy of Management Review, 9*(3), 428-437.

Bach, G. R. & Wyden, P. (1968). *The Intimate Enemy: How To Fight Fair In Love And Marriage.* New York: Avon Books.

Bass, B. M. (1990). The meaning of leadership. In J. T. Wren (Ed), *The Leader's Companion: Insights On Leadership Through The Ages* (37-38). New York: The Free Press.

Balcazar, F. E., Hopkins, B. L., & Suarez, Y. (1985). A critical, objective review of performance feedback. *Journal of Organizational Behavior Management*, *7*(3–4), 65–89.

Berger, C. R. & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: toward a developmental theory of interpersonal communication. *Human Communication Research, 1*(2), 99-112.

Berrett, D. (2012). Too Many A's? Proposal Could Make Them Rarer at U. of Minnesota. *The Chronicle of Higher Education*, (38).

Blumer, H. (1969). *Symbolic Interactionism: Perspective And Method.* Englewood Cliffs: Prentice-Hall.

Braithwaite, D. O. & Baxter, L. A. (Eds.).(2008). *Engaging Theories In Interpersonal Communication: Multiple Perspectives.* London: SAGE.

Buller, D. B. & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory 6*(3), 203-242.

Butler, T. L. (2017). *An investigation of low performing central florida teacher evaluation feedback and improvement plans as related to value-added model scores and instructional practice scores (*Doctoral dissertation). Retrieved from the University of Central Florida library database.

Cairns, J. & Sears, A. (2010). *A Good Book, In Theory: Making Sense Through Inquiry.* Toronto: University of Toronto Press.

Cameron, J., & Pierce, W. (1994). Reinforcement, reward, and intrinsic motivation: a metaanalysis. *Review Of Educational Research*, *64*(3), 363–423.

Carless, D. (2006). Differing perceptions in the feedback process. *Studies In Higher Education*, *31*(2), 219–233.

Castetter, W. B. (1981). *The Personnel Function In Educational Administration.* New York: Macmillan.

"Conflict Theory" (2018). *Investopedia.* Retrieved from: https://www.investopedia.com/terms/c/conflict-theory.asp

Conrad, C. R. & Poole, M. S. (2012) *Strategic Organizational Communication: In A Global Economy* (7th ed.). Malden, MA: Wiley-Blackwell.

Cronin, T. E. (1984). Thinking and learning about leadership. In J. T. Wren (Ed), *The Leader's Companion: Insights On Leadership Through The Ages* (27-32). New York: The Free Press.

Dainton, M. & Zelley, E. D. (2014). *Applying Communication Theory For Professional Life: A Practical Introduction* (2nd Ed). Thousand Oaks, CA: Sage Publications.

Darling-Hammond, L. & Wise, A. E. (1984). Teacher evaluation and teacher professionalism. *Educational Leadership*, *42*, 28–33.

Deetz, S. A. (1994). Future of the discipline: the challenges, the research, and the social contribution. *Communication Yearbook,* 17, 565-600.

DeNisi, A. S. & Kluger, A. N. (2000). Feedback effectiveness: can 360-degree appraisals be improved? *The Academy Of Management Executive,* 14(1), 129-139.

Donaldson, M.L. (2009). So long, lake wobegon? Using teacher evaluation to raise teacher quality. *Center for American Progress.* Retrieved from: https://www.americanprogress.org/issues/education/reports/2009/06/25/6243/so-long-lake-wobegon/

Donaldson, M.L. (2010). No more valentines. *Educational Leadership, 67*(8), 54-58.

Duckworth, A. (2016). *Grit: The Power Of Passion And Perseverence.* New York: Scribner.

Dweck, C. S. (2006). *Mindset: The New Psychology Of Success.* New York: Ballantine Books.

Dziak, M. (2016). Interpersonal communication. *Salem Press Encyclopedia*.

Ehrgott, R. H., Henderson-Sparks, J. C., & Sparks, R. K. (1995). Managing your marginal teachers. *Principal, 74*, 32-35.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363-406.

Fedor, D. B., & Buckley, M. R. (1987). Providing Feedback to Organizational Members: A

    Reconsideration. *Journal of Business and Psychology*, *2*(2), 171–181.

Florida Statute § 1012.33.(2016)

Florida Statute § 1012.34. (2015)

Fuhr, D. L., 1934-. (1993a). Rescuing students from borderline teachers. *Education Digest*, *59*,

    42–44.

Griffin, E. M. (1997). *A First Look At Communication Theory (3ʳᵈ Ed).* New York: McGraw-Hill.

Getzels, J. W. & Guba, E. G., author. (1957). Social behavior and the administrative process. *The*

    School Review, (4), 423.

Haefele, D. L. (1993). Evaluating teachers: A call for change. *Journal of Personnel Evaluation in*

    Education, *7*(1), 21–31.

Hain, J. H. & Smith, G. J. (1996). Evaluation of teachers—the principal's dilemma. Retrieved

    from: https://eric.ed.gov/?id=ED013490

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*,

    77(1), 81-112.

Hattie, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to*

    Achievement. New York: Routledge.

Hayes, R. M., & Schaefer, S. (2009). CEO pay and the Lake Wobegon Effect. *Journal of*

    Financial Economics, *94*(2), 280–290.

Hughes, R. L., Ginnett, R. C., & Curphy, G. R. (1993). What is leadership? In J. T. Wren (Ed),

    The Leader's Companion: Insights On Leadership Through The Ages (39-43). New York:

    The Free Press.

Jerald, C. (2012). *Leading for Effective Teaching: How School Systems Can Support Principal Success.* Retrieved from: https://www.k-12leadership.org/sites/default/files/jerald-white-paper-leading-for-effective-teaching_1_0.pdf

Kane, T. J., Staiger, D. O. (2012). *Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains. Policy and Practice Brief. MET Project*. Bill & Melinda Gates Foundation.

Kane, T. J., Staiger, D. O., Rockoff, J. E., & Taylor, E. S. (2012). Information and employee evaluation: evidence from a randomized intervention in public schools. *The American Economic Review*, (7), 3184.

Keillor, G. (2001, August 26). "In Search of Lake Wobegon." *The New York Times*. Retrieved from https://www.nytimes.com/2001/08/26/books/chapters/in-search-of-lake-wobegon.html

Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research*, (2), 211.

Kruger, J. (1999). Lake Wobegon be gone! The "below-average effect" and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology*, 77(2), 221–232.

Langlois, D. E., & Colarusso, M. R. (1988). Improving teacher evaluation. *Education Digest*, *54*, 13–15.

LaRossa, R., & Reitzes, D. C. (1993). Symbolic interactionism and family studies. In P. G. Boss, W. J. Doherty, R. LaRossa, W. R. Schumm, & S. K. Steinmetz (Eds.), *Sourcebook of family theories and methods: A contextual approach* (pp. 135-166).

Ling, T. W. (2014). *An examination of school principals' moral reasoning and decision-making along the principalship track and across years of experience* (Doctoral dissertation). Retrieved from:

http://etd.fcla.edu.ezproxy.net.ucf.edu/CF/CFE0005368/T_Ling_Diss_Final_Submission.pdf

Livingston, P. & Livingston, D. (2012). *Teacher Development Toolkit: For The Marzano Teacher Evaluation Model.* Prepared for Marzano Research Laboratory. Retrieved from:

http://www.k12.wa.us/TPEP/Frameworks/Marzano/Marzano_Teacher_Development_Toolkit.pdf

Lunenburg, F. C. & Irby, B. J. (2008). *Writing A Successful Thesis Or Dissertation: Tips And Strategies For Students In The Social And Behavioral Sciences.* Thousand Oaks, CA: Corwin Press.

Manasa, K. L., & Reddy, N. (2009). Role of training in improving performance. IUP Journal of Soft Skills, 3(3/4), 72-80.

Mannion, J. (Ed.).(2002). *Essentials of Philosophy.* New York: Fall River Press.

Marzano, R. J. (2007). *The Art And Science Of Teaching: A Comprehensive Framework For Effective Instruction.* Alexandria, VA: ASCD.

Marzano, R. J. (2011). Research Base and Validation Studies on the Marzano Evaluation Model. Retrieved from:

http://www.marzanoevaluation.com/files/Research_Base_and_Validation_Studies_Marzano_Evaluation_Model.pdf

Mela, C. A, (2013). *An analysis of the school board of Brevard county instructional personnel performance appraisal instrument and the validity of its components* (Doctoral dissertation).

Retrieved from:

http://etd.fcla.edu.ezproxy.net.ucf.edu/CF/CFE0004719/Entire_Dissertation_Done.pdf

Moore, J. (2017). Where Is the Critical Empirical Interpersonal Communication Research? A

Roadmap for Future Inquiry into Discourse and Power. *Communication Theory*, *27*(1), 1–20.

Manz, C. C., & Sims, H. P. (1991) Superleadership: Beyond the myth of heroic leadership. In J. T.

Wren (Ed), *The Leader's Companion: Insights On Leadership Through The Ages* (212-221).

New York: The Free Press.

Myers, D. G. (2000). The psychology of humility. In R. L. Herrmann (Ed), *God, Science, and*

Humility: Ten Scientists Consider Humility Theology (153-175). West Conshohocken, PA:

Templeton Press.

Office of the Governor, State of Georgia, Special Investigators. (2011). Report to the Governor on

Cheating in the Atlanta Public School System. Atlanta, GA: Georgia

Ovando, M. N. (1992). *Constructive Feedback: A Key to Successful Teaching and Learning*.

Ovando, M. N. (2005). Building Instructional Leaders' Capacity to Deliver Constructive

Feedback to Teachers. *Journal of Personnel Evaluation in Education*, *18*(3), 171–183.

Owens, R. G., & Valesky, T. C. (2015). Organizational behavior in education: Leadership and

school reform (11th ed.). Boston, MA: Pearson.

Pace, D. P. (2015). *A study of the impact of Brevard public school's peer coaching model on*

student achievement outcomes and teacher evaluation results (Doctoral dissertation).

Retrieved from: http://etd.fcla.edu.ezproxy.net.ucf.edu/CF/CFE0005678/Pace_Final_pdf_2-

16-15.pdf

Pearce, W. B., & Cronen, V. (1980). *Communication, action, and meaning: The creation of social realities*. New York: Praeger.

Pearce, W. B., & Cronen, V. (1982). The Coordinated Management of Meaning: A theory of communication. In F. E. X. Dance (Ed.)., *Human communication theory, 61-89.* New York: Harper & Row.

Phillips, M. A. (2014). *An analysis of the variables in implementation of the Marzano teacher evaluation system in the state of Florida* (Doctoral dissertation). Retrieved from: http://etd.fcla.edu.ezproxy.net.ucf.edu/CF/CFE0005399/Phillips_dissertation_pdf_7-04-14.pdf

Pink, D. H. (2011). *Drive: The Suprising Truth About What Motivates Us.* New York: Riverhead.

Pryor, K. (1999). *Don't Shoot The Dog! The New Art Of Teaching And Training* (Revised Ed). New York: Bantam.

Rafalski, S. H. (2015). *Policy implications of a teacher evaluation system: the relationship of classroom observations, levels of feedback, and student achievement outcomes* (Doctoral dissertation). Retrieved from: http://etd.fcla.edu.ezproxy.net.ucf.edu/CF/CFE0005693/Rafalski_final_pdf_2-16-15.pdf

Reiner, R. (Producer & Director). (1995). *The American President* [Motion picture]. United States & Canada: Columbia Pictures.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, (2), 119.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189.

Slaughter, J. B. (2012). Thoughts on educational leadership. In K. S. Gallagher, R. Goodyear, D. J. Brewer, and R. Rueda (Eds) *Urban Education: A Model For Leadership And Policy.* New York: Routledge.

"Statistical Power Analysis." (2018). *Statistics Solutions: Advancement Through Clarity.* Retrieved from: http://www.statisticssolutions.com/statistical-power-analysis/

Steinberg, W. J. (2011). *Statistics Alive!* (2nd Ed.). Washington DC: Sage.

Stewart, J. R. (1999). *Bridges Not Walls: A Book About Interpersonal Communication.* New York: McGraw Hill.

Strenth, R. S. (2013). *The moral reasoning and moral decision making of urban high-poverty elementary school principals in a large urban southeastern school district* (Doctoral dissertation). Retrieved from: http://etd.fcla.edu.ezproxy.net.ucf.edu/CF/CFE0004943/Strenth_Robert_S_201305_EdD_2.pdf

Stronge, J. H. (1995). Balancing individual and institutional goals in educational personnel evaluation: a conceptual framework. *Studies in Educational Evaluation*, *21*(2), 131–151.

Tucker, P. D. (1997). Lake Wobegon: where all the teachers are competent (or, have we come to terms with the problem of incompetent teachers?). *Journal of Personnel Evaluation in Education, 11,* 103-126.

Tyler, J. H., & Taylor, E. S. (2011). *The Effect of Evaluation on Performance: Evidence from Longitudinal Student Achievement Data of Mid-Career Teachers.* Retrieved from: http://www.nber.org.ezproxy.net.ucf.edu/papers/w16877

Vitton, C. J., & Wasonga, T. A. (2009). Between Kohlberg and Gilligan: levels of moral judgment among elementary school principals. *Leadership & Policy in Schools 8*(1), 92-116.

Watzlawick, P., Weakland, J. H., & Fisch, R. (1974). *Change: Principles of Problem Formation and Problem Resolution.* New York, Norton.

Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., & New Teacher Project. (2009). *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness. Executive Summary. Second Edition*. New Teacher Project.

Wiese, D. S. & Buckley, M. R. (1998). The evolution of the performance appraisal process. *Journal of Management History*, (3), 233.

Wheeler, P., & Haertel, G. D. (1993). Resource handbook on performance assessment and measurement.

Wolf, J., & Wolf, K. (2013). The Lake Wobegon Effect: Are All Cancer Patients above Average? *MILBANK QUARTERLY*, *91*(4), 690–728.