2-21-2023

Impacting African American Student Achievement in the Middle School STEM Classroom by Teaching Mathematics Through Arts Integration and Design Thinking

Robert Lippert
University of Missouri-St. Louis, dr.rob.lippert@gmail.com

Tjuannia R. Seals
University of Missouri-St. Louis, trs018@umsl.edu

Impacting African American Student Achievement in the Middle School STEM Classroom by Teaching Mathematics Through Arts Integration and Design Thinking

Robert J. Lippert
M.Arch., Architecture, Washington University St. Louis, 1996
B.T. Architectural Technology, Washington University St. Louis, 1994

Tjuannia R. Seals
MA, Educational Technology, Lindenwood University, 2015
M.Ed., Educational Administration, University of Missouri-St. Louis, 2000
BS, Education, Harris-Stowe State University, 1996

A Co-Authored Dissertation submitted to
The Graduate School at the University of Missouri-St. Louis
in partial fulfillment of the requirements for the degree
Doctor of Education with an emphasis in Educational Practice

May 2023

<u>Dissertation Committee</u>

Helene J. Sherman, Ed.D.
Chairperson

Charles R. Granger, Ph.D.

Keith W. Miller, Ph.D.

Abstract

Middle school is a transitional period in which many students experience content-specific teachers, travel between classrooms, and explore extracurricular activity options for the first time. Historically, African American middle school students have not fared well in Science, Technology, Engineering, and Mathematics (STEM) on standardized assessments, performing significantly below their Caucasian counterparts in mathematics and science. From the beginning of their academic careers, a lack of access to quality teachers, excessive use of direct instruction strategies, and a lack of resources in their school communities, contribute to their overall apathy towards the subject matter and factor into their underperformance. As a result, fewer African Americans pursue STEM studies in secondary education resulting in underrepresentation in STEM-related professions.

To stimulate African American students' interest in the field, the approach to STEM instruction requires alternative strategies. Some students who do not effectively demonstrate and communicate their understanding of mathematics or science principles using traditional equations instead show a clearer expression of knowledge through alternative, more artistic media. The extent of the effectiveness in implementing a design thinking and arts integrated project-based learning activity (DAIP) to increase African American students' interest and achievement in STEM subjects was explored.

Acknowledgements

Many thanks to Stacey Miller who trusted us and followed our lead as we used her classroom as a laboratory. We would like to thank the middle school students and their parents who participated in this study. We appreciate their hard work and willingness to be creative and try something new. A special thanks to Ms. Rachel Clark for her support in allowing us to conduct the research in her school.

Thank you to our advisors, Dr. Helene Sherman, Dr. Charles Granger, and Dr. Keith Miller, for sharing their knowledge, experience, and continued encouragement. Thanks is extended to all of the members of the UMSL STEM THEMES cohort for their suggestions and assistance throughout the doctoral program. Thank you to Dr. Dean Coble for sharing his statistics expertise as we worked out how to analyze this data.

Finally, we owe a debt of gratitude to all of our teaching colleagues for their understanding, encouragement, and cooperation as we took on this monumental task.

Dedications

Author 1: Robert J. Lippert

I would like to dedicate this dissertation to my wife, Dr. Kathleen Dwyer, for her continuous love and support. You worked extra as hard and made many sacrifices to empower me to earn this degree. Also, to my parents for instilling in me a work ethic, especially my mother, Rita for her unwavering encouragement. To my co-author, Tjuannia, we made a great team! Special thanks to Dr. Matt Troutman, Ms. Jane Roth and Ms. Julie Northrip for flexibility with work commitments while I went through this 3-year journey.

Author 2: Tjuannia Seals

I would like to dedicate this dissertation to my mother, Zelma Myers, who is my biggest cheerleader. You have supported me throughout my entire education journey and I love and appreciate you for it. To my “crew” Cheryl Howard, Leontyne Miller, Valerie Morris, and Crystal Spearman who always go above and beyond to encourage and uplift me. Undeniably, to my co-author Rob who is the BEST writing partner ever. You have truly made this an enjoyable experience. I would also like to extend a big thank you to my family for always being so reassuring and believing in me.

Table of Contents

List of Tables

List of Figures

Chapter 1
Introduction

Introduction

Educational opportunities for African Americans before emancipation were extremely limited (Darling-Hammond, 2000). Even with the elimination of slavery in 1865, African Americans in the United States have experienced disparities in the educational school system ranging from funding inequities, lack of qualified teachers and limited resources, to outright oppression and denial of educational liberties. The 1896 court ruling of Plessy vs. Ferguson (Plessy v. Ferguson 1896, 2022), the institution of Jim Crow laws, and redlining housing locations have each significantly impacted and limited the potential success of some African American students, both in the past and today (Anderson, 2017). According to Allen and Jewell (1995), opportunities for this marginalized group to have equal access to education have been beset with barriers from the very beginning. Racism and restricted freedoms experienced by African Americans have led to an unbalanced educational experience. Drastically different learning opportunities are strongly related to vast inequities in student achievement (Darling-Hammond, 2000). Many urban and suburban school systems have a disparity in academic resources available to their respective staff and students due to the community in which they reside (Jenkins, 2017).

The disparities in educational opportunities and resources over time have translated into substandard performance by African American students in underserved schools, deflated the confidence of minority students in their ability to succeed in Science, Technology, Engineering and Mathematics (STEM) subject matter and careers, and resulting in a disproportionately fewer number of these students succeeding in STEM topics or entering STEM careers. Changing this self-defeating attitude requires more effective instruction for these students in STEM topics. Accomplishing this in underfunded schools requires equalizing the resource disparities through novel teaching strategies rather than relying on expensive investments that are unlikely to appear in the foreseeable future.

Background

Research indicates that many African American children demonstrate lower reading and mathematics scores when entering school. These deficiencies can lead to underperformance throughout their academic careers (Dahl & Lochner, 2005; V. E. Lee & Burkham, 2002). For many African American students, access to a rigorous STEM education is challenging (Ta n et. al 2018). Critical-thinking and problem-solving skills exercised through hands-on participation, frequent writing, and careful reading of texts have not been the focus for many urban schools. These schools tend to focus more on rote learning. (Cooper & Sherk, 1989; Darling-Hammond, 1997). Many African American student challenges in STEM topics have begun by the time they reach the middle school level (Solomon, 2013). The foundational skills and concepts acquired in earlier grades are not often reflected in student comprehension as they encounter more complex material. As a result, their ability to transfer knowledge and apply processes to new situations is inhibited. According to the U.S. Department of Education, “African-Americans received just 7.6 percent of all STEM bachelor's degrees and 4.5 percent of doctorates in STEM.” (USDOE, 2016, p. 2).

The reinforcement of STEM concepts is most effective when using hands-on, project-based learning activities that incorporate real-life situations and place an emphasis on critical thinking and collaboration (Jenkins, 2017). By design, integrating STEM topics into a well-planned project-based learning (PBL) lesson/activity facilitates learning where students gain a holistic understanding of how STEM subjects are interconnected. (Chine, 2022). Instead of embracing STEM subjects, most African American students at the middle school level choose to focus on arts-related or elective courses including art, choir, music appreciation, drama, and band (Brazee, 2000). Of this student population, a great majority achieve in these classes when they are allowed to creatively express themselves. As an antidote, the employment of an iterative and non-linear problem-solving approach may positively impact student performance in STEM disciplines by African American students.

Problem

Few African American students in underserved schools demonstrate optimum success in STEM classrooms where traditional methods, such as direct instruction, are implemented (Gonzalez & Kuenzi, 2012). These students often disengage from the topics and feel incapable of succeeding in STEM fields (Johnson & Kritsonis, 2006). Evidence of the efficacy of Jerome Bruner's theory of discovery learning (Bruner, 1961) indicates that students who build knowledge in a guided manner through uncovering evidence and making connections, retain this information and can apply it in creative ways, thereby achieving a higher level of learning. By exploring different instructional methods in middle school, it seems possible that alternative problem-solving techniques may successfully stimulate learning in some African American students.

Additionally, not all students effectively demonstrate and communicate their understanding of mathematics or science principles using traditional equation based instruction. Instead, they more clearly express knowledge through alternative, more artistic media (Jenkins, 2017). Using design thinking to solve a project-based activity and then explaining their understanding using the arts may be a means of capturing and retaining African American student interest in STEM subjects. The iterative problem-solving process of design thinking has evolved beyond a planning tool used by teachers to prepare lessons. This method of problem-solving has begun making its way directly into classrooms for a number of years, yet little research exists on the effectiveness of teaching and employing design thinking with students younger than at the secondary education level (Pande & Bharathi, 2020).

Purpose

The integration of arts pedagogy using design thinking connected to problem-solving strategies may effectively equip students for success in the STEM classroom. Design inquiry allows teachers to use their expertise to create targeted learning experiences intended for students to develop stronger critical thinking skills. Through project-based learning, students learn cross-disciplinary skills by developing solutions to real-world problems. Cross curricular planning and dialogue enables teachers to structure lessons to meet the needs of students of varying skill levels. The validation of the effectiveness of this alternative strategy was explored. This instructional strategy was analyzed to determine its effect on African American middle school students' attitudes and confidence toward learning STEM content, specifically mathematics. By determining the extent to which the effectiveness in implementing a design thinking and arts integrated project-based learning activity (DAIP) increases STEM competency, teachers in underserved schools may be equipped with a new teaching tool to address the disparities experienced by African American students.

Mathematics was chosen to specifically measure STEM improvement using DAIP because mathematics is the only discipline universally taught at the middle school level and required by state standards. Being taught to all students, mathematics proficiency is readily recorded through existing assessment techniques and provides a reliable basis for comparison and control. Science, technology, and engineering all appear in middle school curricula to varying degrees, but are not consistently assessed in middle school.

Research Questions

1. To what extent would minority students' STEM scores in a disenfranchised community improve when taught STEM concepts using an alternative problem-solving teaching methodology (DAIP)?
2. To what degree would minority students in a disenfranchised community improve in confidence and interest in STEM concepts when taught mathematics using an alternative problem-solving strategy (DAIP)?

Hypotheses

$H_0$1: There is no significant difference in the achievement in pre-algebra as measured by the Renaissance STAR Benchmark Assessment (Renaissance, 1998) and a teacher-generated pre- and post-assessment between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

H_a1: There is a significant difference in the achievement in pre-algebra as measured by the Renaissance STAR Benchmark Assessment (Renaissance, 1998) and a teacher-generated pre- and post-assessment between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

H_02: There is no significant difference between the confidence and interest in STEM topics as measured by scores on the Attitudes Toward Mathematics Inventory (ATMI) (Tapia, 1996) between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

H_a2: There is a significant difference between the confidence and interest in STEM topics as measured by scores on the Attitudes Toward Mathematics Inventory (ATMI) (Tapia, 1996) between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

Significance

The use of DAIP could be paramount in changing attitudes and broadening critical thinking and problem-solving skills in a non-restrictive manner. Historically, African American students in urban schools have lacked resources and funding afforded to their Caucasian counterparts in wealthy, suburban districts and consequently do not have the benefit of learning using the latest tools or acquiring information from the most qualified educators. The result is often that their teachers are forced to deliver direct instruction with minimal active learning, leaving students struggling to engage in particularly challenging subjects and feeling disconnected from STEM topics.

Teaching through the framework of a real-world problem that is relevant to the students, project-based learning has the potential to provide teachers in these underserved schools with a relatively low cost alternative to direct instruction. In this model, students actively design a solution to that real-world problem while gaining knowledge about STEM subjects and exercising STEM skills. By allowing the students to express their new understanding in creative ways through art, African American students can connect with science and mathematics and begin to recognize that they can aspire to STEM careers. Once students are taught to think critically, the intended goal is for that ability to be applied in all facets of their personal, educational, and professional journeys.

Knowledge gained through testing the effect of design thinking can potentially provide a pathway for opportunities leading to expanded teaching strategies in the STEM classroom and provide alternative methods for instilling confidence in science and mathematics concepts in middle school students.

Definition of Terms

Disparity of African American Students' Achievement in STEM Education

The underrepresentation and underperformance of African Americans in science, technology, engineering, and mathematics (National Science Foundation, 2001) in schools and the workforce.

Project-based Learning

Project-based learning (PBL) is a classroom instructional strategy designed to engage students in investigating real-world problems in an authentic, relevant context using inquiry through the execution of a project activity. This strategy differs from explicit instruction in that the teacher facilitates students to engage in active learning as opposed to the teacher directly delivering material. Skills and knowledge in the subject are acquired by practicing while testing their proposed solutions to the real-world problem.

Design Thinking

A problem-solving approach with a focus on the result. It incorporates a 5-stage thinking process calling for one to "empathize, define, ideate, prototype, and test" (Stanford, 2010, pp. 1-5) in an iterative process, developing a solution that is refined with each subsequent prototype and test experience.

Arts Integration

A teaching strategy where students construct an understanding and communicate that understanding through an art form. Employing the creative process, students connect the art with knowledge from other disciplines, achieving competencies in both subjects simultaneously (Silverstein & Layne, 2010).

DAIP

Design arts integrated project (DAIP) is a teaching strategy, based upon project-based learning, which uses design thinking to solve a real-world problem, expressing acquired knowledge with an arts component.

Math-A-Thon

A Math-A-Thon is a school sponsored charity fundraiser allowing students to secure pledges and donations for solved math problems.

Constructivism

An educational theory based on the idea that people actively construct, or make their own knowledge, and participate in constructing their learning. Learners use prior knowledge to build upon new ideas, a system referred to as a spiral curriculum (Bruner, 1960). This is more effective when it happens collaboratively, in a social environment (Vygotsky, 1978).

Discovery Learning Theory

The theory underlying inquiry-based learning (Bruner, 1961) is where learners build on prior experiences and knowledge. With roots in Constructivism (Bruner, 1960) the learner uses their intuition, imagination, and creativity to discover additional facts and correlations which add to their knowledge.

Explicit Instruction

A methodological teacher centered strategy employed for delivering specific concepts and procedures in a particularly formatted and carefully sequenced manner, directed and focused by the teacher.

Assumptions

African American students, by the nature of their social, economical, and cultural status, are often at a disadvantage with regards to education in general and specifically with STEM education. Decades of marginalization by a socio-economic system that favors a Caucasian majority has resulted in an underrepresentation of African American individuals in STEM fields. As a result of generations of exclusion in STEM fields, a culture of non-success in STEM has driven African American students away from these subjects.

For practical implementation, the intervention focused on eight pre-algebra classes at a single middle school in a district that is almost exclusively African American. Four sections were taught using a project-based learning activity, expressed through an artform. The remaining four sections were taught the concepts using traditional empirical lessons with direct, explicit instruction. With the intent to limit the impact on the time involvement of both the teacher and the students, the intervention centered on a single, eight-week unit.

The sample selection was not random since the students were assigned by section by the school, resulting in a quasi-experimental design (Creswell, 2014). Each pre-algebra section was randomly assigned to one of two groups; Treatment or Control, and all students in both groups received the same pre- and post-assessments and surveys.

The results represent only a single data set extracted from one curricular unit taught in eighth grade with a sample group of students from one school. Because the sample pool was modest and the duration of the study short, clear determination of the efficacy of the teaching strategy was often difficult. Additionally, design thinking as a problem-solving technique, like many instructional strategies, takes practice to be executed well.

Summary

While redistributing financial resources and providing additional funding to lower-income schools could quickly impact student success, that solution is mired with political ramifications. An alternative approach is to alter the quality and type of teaching strategies in underfunded classrooms using materials with minimal cost. Implementing DAIP in STEM classrooms has the potential of equalizing the disparities in resources that middle school students in African American neighborhoods face. With its theoretical roots in the spiral curriculum concept of constructivism (Bruner, 1961), design thinking implemented as a problem-solving technique for these activities provides students with a less linear, less formally structured environment incorporating minimal direct or explicit instruction. By tapping into the expressive nature of these students and allowing them to communicate their understanding through the arts, disenfranchised African American middle school students may be equipped with critical thinking and problem-solving skills critical for success in STEM studies. Academic success in the classroom may then foster interest in entering STEM career fields.

While some anecdotal evidence exists in the literature to suggest that this model shows promise, implementation has been primarily at the secondary education level. To date, little research had been done which quantified the effectiveness with the middle school demographic and virtually none with African American students. The intention was to test the efficacy of design thinking and arts integration in a project-based learning environment, quantifying potential increases in student success and measuring improved attitude toward STEM in disenfranchised African American middle school students.

Chapter 2
Review of the Literature

Introduction

Chapter 2 explores the factors that underlie the challenges African American middle school students often face when studying STEM topics. Challenges can affect students' participation and interest in STEM areas of study. A contributing factor to these learning barriers may be a lack of student confidence induced by the rigidity inherent in direct instruction often employed in teaching STEM (Museus et al., 2011; Turner et al., 2010). Minority students may lean toward the arts because of the expressive freedom and creativity that arts can offer (Brazee, 2000). Growing evidence indicates that project-based learning provides an expressive and dynamic alternative to explicit instruction for African American students (Jenkins, 2017).

A project-based activity with an arts component and utilizing design thinking as the problem-solving method may provide a means of engaging a segment of the student population feeling disenfranchised from STEM. The literature review explores the background and effectiveness of project-based learning, design thinking, and arts integration as instructional tools with specific examples of their application to STEM education and concludes with an examination of constructivist and discovery learning theories as a research lens to frame a design-thinking based project.

Search Description

A broad search was conducted for recent literature using online databases. Primary databases queried were ERIC and JSTOR, with limited use of EBSOhost, Education Full Text, and SCOPUS. Additionally, online searches were performed on Google and Google Scholar. Primary keywords in the search included: *design thinking, African American student STEM achievement, middle school STEM achievement, arts integration, constructivism,* and *discovery learning*. Resources were evaluated for relevance to the research questions and summarized to identify key findings and underlying frameworks that may inform the pursuit of the research questions. Additional sources were also identified from citations and the bibliography included in relevant articles.

African American Middle School Student Disparity in STEM

Adolescence represents a developmental phase of diminishing academic motivation in STEM domains. Middle school is a critical time period to foster student interest and readiness in STEM subjects and future careers (Dantley & Leonard, 2010). Unfortunately, not all middle school students are given the latitude "to engage, learn, and achieve in STEM subject areas" (Morena et al., 2016, p. 1). For many of these students, access to relevant and inspiring experiences in STEM is limited, thereby resulting in a lack of representation in future STEM careers (Hannemann, 2007). Introducing STEM topics at the elementary and middle school levels can help to increase student interest and academic performance as well as inspire them to pursue advanced studies in STEM related career fields (Dabney et al., 2012; Tai et al., 2006). Quite often when some students enter middle school, interest in STEM subjects is lost and they fail to be successful in those areas (Museus et al., 2011; Turner et al., 2010). Their experience or inexperience in middle school leads to their inability to be successful at the high school and postsecondary levels which, in turn, makes them unwilling and unable to pursue STEM studies and careers.

According to the White House Initiative on Educational Excellence for African Americans, this demographic of students from disenfranchised communities lacking resources commonly display substandard academic performance and tend to not envision themselves in STEM fields (U.S. Department of Education, 2016). African American underperformance in the STEM areas of mathematics and science is a national problem. From primary to secondary education in mathematics and science, quality instruction has been scarce for students of color (Dantley & Leonard, 2010). African American students have consistently underperformed in comparison to their white peers in the STEM areas of mathematics and science.

Specifically, African American students underperform in mathematics compared to their Caucasian counterparts (National Center for Educational Statistics, 2019). Disparities in the resources provided to African American students in schools in the United States translate into real differences in performance. The resulting problem of “making American schools adequate learning institutions for all students is an on-going challenge” (Johnson & Kritsonis, 2006).

In addition to limited resources, poor instruction, unchallenging work, and tracking into low-level classes, a primary reason for student disaffection and poor performance in mathematics has been the disconnect between the curriculum and their cultural orientation (Dantley & Leonard, 2010). Gloria Ladson-Billings (1994) suggests that by empowering students intellectually and socially, while using references to their culture, may improve learning and attitudes; bridging the disconnect. Project-based learning can become the tool that will allow students to work together and engage in activities that will enable them to solve relevant, real-world problems in STEM areas.

Project-based Learning

Defining Project-based Learning: Project-based learning is characterized by self-directed learning and extended engagement with learners. This method of instruction puts learning in a real-life situation and places heavy emphasis on reasoning strategies and collaboration (Buck Institute for Education, 2014). Project-based learning (PBL) is a classroom instructional strategy designed to engage students in investigating real-world problems in a relevant context incorporating inquiry (Cervantes et al., 2015). PBL assists learners with the development of creative-thinking, problem-solving, and communication skills while collaboratively solving multi-faceted problems through creating projects (Bell, 2010; Duch et al., 2001; Hmelo-Silver, 2004; Jonassen, 2000, 2004; Savery, 2006; Sendag & Ferhan Odabasi, 2009). Students who set goals and tasks and monitor their own progress independent of their teachers are more likely to achieve a higher level of success (Sunger et. al., 2006). Project-based learning is based on constructivist principles. Utilizing collaboration and personal autonomy, students are mentored by adults. (Egenrieder, 2010). Krajcik, Blumenfeld, Marx, and Soloway (1994) and Thomas (2000), describe a productive project-based learning atmosphere consists of the following five components for authenticity:

(1) an engaging question,

(2) student-created products,

(3) collaboration,

(4) an audience, and

(e) technological based cognitive and communication tools.

Project-based Learning in the Classroom: The purpose of project-based learning is to engage students in actual investigative problem-solving techniques to deepen learning and understanding of complex problems. (Blumenfeld et al., 1991; MaKinster et al., 2001; McGrath, 2004). The principles of PBL allow students to ask and refine questions, make predictions, defend ideas, develop plans, gather and interpret data, draw conclusions, articulate ideas, and create artifacts (Blumenfeld et al., 1991; Mergendoller et al., 2006; Thomas, 2000). Students experience authentic problem-solving situations that allow them to make connections between classroom theory and real-life occurrences (Blumenfeld et al., 1991). The students are tasked with exploring problems and working cooperatively with their fellow classmates and community resources (MaKinster et al., 2001). Students involved in PBL classrooms tend to showcase a higher level of motivation than those individuals who do not engage in a PBL experience (Blumenfeld et al., 1991). Emotional and intellectual stimulation can lead to academic achievement in the classroom (Allen et. al 2013). The three goals associated with teachers increasing student motivation include deliberately involving students in the classroom, encouraging students to engage cognitively in the learning process, and helping students to become intrinsically invested in lifelong learning. Student interest is paramount and required for project-based learning to have authenticity. In addition to students working collaboratively in making decisions on every facet of the project from start to finish, the ability for students to make their own assumptions on how to complete a project gives them a vested interest in the results (Robinson, 2013).

Project-based Learning and STEM Education: STEM PBL is rooted in the theory of constructivism calling for students to engage in problem-solving, interdisciplinary instruction, constructed responses, hands-on, and interactive group activities (Capraro & Slough, 2008; Clark & Ernst, 2007; Dolmans et al., 2005). Historically, students lose interest in STEM topics once they experience failure in their academic performance. Project-based learning can generate renewed interest in STEM subjects and STEM related careers (Egenrieder, 2010). PBL consists of student-centered performance activities, formative assessments, and community-based learning environments. Previous studies have indicated that lower performing students can be motivated using PBL in comparison to high achievers (Horan et al., 1996). The structure and process of Project-based learning applied to STEM is unique because it addresses and incorporates multiple disciplines within the solution of the project. It begins with indicating explicit outcomes, targeting objectives, and detailing the final project assessment (Hanif et al. 2019).

Implementing Project-based Learning: PBL requires teachers to be more flexible in their classroom by tolerating a higher-than-normal noise level as students collaboratively engage in activities. Teachers must endorse the role of facilitator to support students in acquiring new skills and believe their students are capable of learning without being explicitly instructed. (Condliffe et al. 2017). Classroom management raises concerns for some teachers in a project-based learning environment as students participate in activities for an extended amount of time and misbehavior and inattentiveness arise (Thomas, 2000 & Hertzog, 2007). It is imperative for teachers to establish, model, and maintain classroom expectations for effective cooperative group activities (Darling et al., 2008).

Arts Integration

Defining Arts Integration: Arts integration is a strategy allowing students to make sense of core curriculum through an art form. It bridges the gaps between abstract and concrete conceptual ideas, moving beyond traditional methods of information acquisition in support of innovation by having students create something new using the academic information that they receive (Marshall, 2014). According to Duma and Silverstein (2014), students partake in the creative process to meet the objectives of both the art and core curriculums. Arts integration can have a particularly positive impact on disenfranchised learners (Duma et. al, 2014). Art design inquiry is grounded in creative thought and problem-solving. It values individual experimentation and encourages the idea that students can use their imaginations to find more than one way to approach, or indeed to solve, a question or problem. Different ways of thinking are encouraged; open-ended questions are asked (Grushka et. al, 2018).

Why Integrate Art in the Classroom: Art education generally focuses on aesthetics and creativity in isolation; not so much the intersection of other subject areas. Because of a lack of experience in arts integration, many educators tend to shy away from using art techniques in core areas (Efland, 2002). Integrating arts in the classroom encourages students to engage in education through a tactile and expressive way (Baird, 2015). Integrating arts in core curricular classes may enhance understanding of difficult concepts. By integrating arts with math, students may be assisted in learning math differently and be provided with a context in which to apply their learning. An art modality included in math content can provide students with more accessibility to the information by allowing students to view math in a nontraditional way (Lau, 2014). Integrating the arts in core subject areas can reinforce specific academic skills and help to improve student reading comprehension (Kabilan & Kamarrudin, 2010; Standly, 2008) and mathematics skills (Moore & Linder, 2012). Arts integration can also improve upon individuals' cognitive development (Hetland & Winner, 2004), as well normed-referenced assessments (College Board, 2013; Whisman & Hixson, 2012).

Art and STEM Education: According to Riley (2013, p. 1), "The Institute for Arts Integration and Steam defines STEAM Education as an approach to learning that uses Science, Technology, Engineering, the Arts and Mathematics as access points for guiding student inquiry, dialogue, and critical thinking." STEAM connects student learning with components of art and core curriculum standards while helping to foster an all-encompassing learning environment that promotes student autonomy and engagement through a holistic learning approach (Lathan, n.d.).

Arts Integration Implementation: Implementing arts integration design techniques in core curriculum will require professional development for content-specific teachers. Professional development in the educational environment is used to strengthen teacher instructional practices for the duration of their career (Mizell, 2010). Effective professional development equips educators with pertinent, relevant skills to meet the academic and emotional needs of all learners. It requires careful planning and implementation with descriptive feedback so educators can swiftly incorporate new practices. Professional development is only effective when teachers use it to improve their instructional practices (Mizell, 2010). Arts integration professional development must enable participants to design targeted activities that engage students in completing complex tasks spanning multiple disciplines.

Design Thinking

Defining design thinking: Ineta Luka (2020) describes design thinking as both a practical process for problem-solving but also a paradigm. In the standard process of deductive problem-solving, we are given the "what" and the "how" of the problem and from these, predict a result. In a similar vein, inductive problem-solving, which is the core principle of the scientific process, assumes that we have a "what" and an observable result and are trying to figure out the "how." With design thinking, we are only given the result and define the "what" and "how" by drawing from experience and prior knowledge. These variables are then tested against the guiding parameters of the desired result to determine if they provide a satisfactory solution (Dorst, 2011). According to the Stanford University Institute of Design, the model assumes that the process can be done by beginners or experienced students alike (Stanford D-School, 2010).

This idea aligns with Jerome Bruner's concept that any material can be taught to any age of student in a way that is appropriate to the student's level of development (Bruner, 1961). It allows students at varied levels to gain knowledge through the process, even if they are experiencing it together simultaneously. "WHY assume a beginner's mindset? We all carry our experiences, understanding, and expertise with us. These aspects of yourself are incredibly valuable assets to bring to the design challenge – but at the right time and with intentionality. Your assumptions may be misconceptions and stereotypes and can restrict the amount of real empathy you can build. Assume a beginner's mindset in order to put aside these biases, so that you can approach a design challenge with fresh eyes" (Stanford D-School, 2010, p. 6). Because they each bring unique experiences and knowledge to the team, beginners and experienced students learn from each other through inquiry.

Design thinking has been used as a model for decades as a method of problem-solving taught to engineering and architecture students (Rowe, 1994). The premise of design thinking differs from traditional *deductive* reasoning where a "what" or object we observe and a "how" or "why" working principle are given and we seek to demonstrate a "result" through observation and testing. In a similar way, *induction* is employed when the "what" and the observed "result" are given and the goal is to determine the underlying principles that are in play causing the behavior that we observe in the "what" (Dorst, 2011, p. 523.) Industrial design professor Kees Dorst (2011, p. 524) introduced a third paradigm, *abduction* in which only the result is given and neither the "what" nor the "how" is defined. Instead of employing pure trial and error to find the desired result which can be an inefficient method, the problem is "framed" with guidelines or principles drawn from previous experience to provide a value for the second variable. The problem can then be tested, observed, and adjustments made to refine the principle and the "what." The revised combination can then be tested, observed, and further adjustments made. This process continues iteratively until the desired "result" is achieved. By applying some previous guidelines drawn from experience, the design thinker narrows the possibilities, focusing on probable solutions, instead of exploring all possible solutions presented by trial and error. Dorst (2011, p. 525) notes that the process begins almost as induction but taken in reverse order.

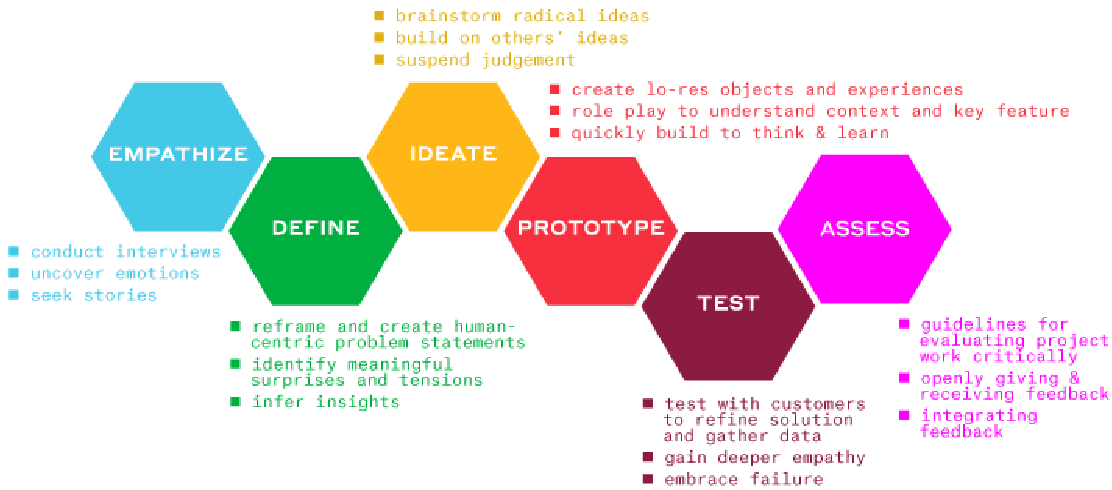
The manual developed by Stanford's design school to teach novice designers how to tackle this problem introduces one more step in the iterative process, *empathy*. This concept originated in the architecture and engineering fields where the designer is ultimately solving a problem presented by someone else, for example, the perfect kitchen, workspace, or the most ergonomic chair. By placing themselves in the position of the person needing the solution and then drawing from their experience allows the designer to choose a baseline working principle to apply to the desired "result" (Stanford, 2010, p.1.) From this baseline, the iterative process can begin as described by Dorst (2011.)

The next step in the process of employing a design thinking instructional strategy is that the designer would define or "frame" the problem by choosing from experience the guideline that most closely aligns with the needed outcomes or objectives. Utilizing this framework, the designer can then begin to develop testable prototypes. At this step, several possible solutions may be presented through brainstorming. These are judged against the result and those most promising are prototyped or prepared for testing.

The possible solutions are tested, comparing them to the desired result and each is documented. As test results are documented, it is important to note the flaws or errors in the solution that prevent that solution from meeting the desired objective. After testing several of the brainstormed solutions, the results are compared, looking for patterns both in the errors and the successes. This iterative process is represented in Figure 2-1.

Figure 2.1

Design Thinking Process

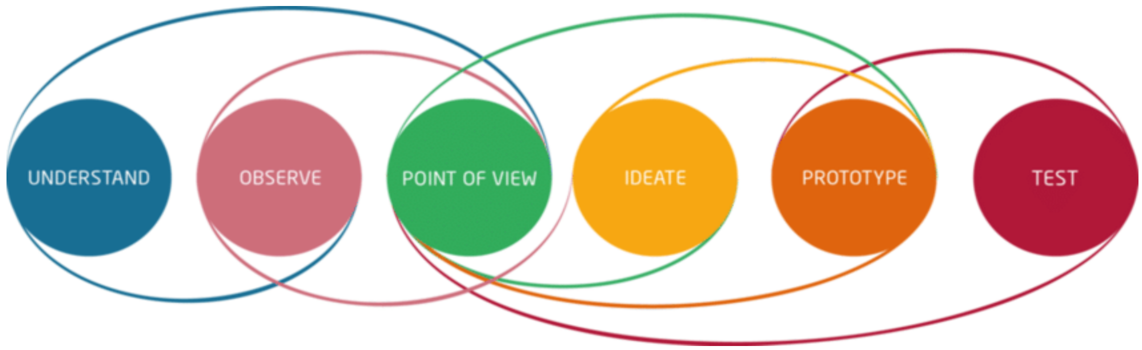


Note: The design thinking process from the Stanford University Institute of Design (2010).

Identifying patterns in the test results helps to refine the guidelines and the next iteration of prototypes are developed. Ranking test results by the degree to which they solved the problem allows the designer, which in this case is the student, to make choices as to which should be developed further and which should be eliminated. The design thinking process is iterated (see Fig. 2-2) until one solution is identified and refined to the degree of accuracy expected.

Figure 2.2

Iterative Design Thinking Process



Note: The iterative nature of the design process (Plattner, et al, 2009, p. 114.)

Design Thinking in the Classroom: With origins in the fine arts, design thinking has been taught for decades as a method of problem-solving for architecture, industrial design, and engineering students at the collegiate level (Pande & Bharathi, 2020). Historically, design has been relegated to the arts and excluded from traditional academic subjects, such as math and the sciences. Instead, creating new concepts and ideas in these fields has been the purview of professionals, not students (Li et al. 2019). Since the 1980's, design thinking has grown to be applied to a broader variety of disciplines. One of the earliest disciplines to adopt this process was business education, using it to teach graduate students to tackle complex marketing strategies when their markets, customers, and products had multiple aspects or variables to be considered. Melles, et al (21012), describe the development and implementation of a design thinking course at Swinburne University. The course was taught simultaneously in Melbourne and Hong Kong and enriched the curriculum of their undergraduate program in management (Melles et al., 2012). Kimball (2011) discusses her experience in teaching design thinking to business management students to prepare them for complex problems faced in practice. Within language arts, the design process is often applied to a non-STEM field: writing. The techniques are the same and demonstrate that for problem-solving, the iterative spiral process where the text becomes more refined and focused with each pass encourages students to recognize opportunities for creative alternatives to traditional thoughts (Leverenz, 2014). Design thinking focuses on having the students define the parameters of the problem while dealing with ambiguity or uncertainty in the target, then actively solve the problem by making connections with the problem to experiences in their lives both in and outside of school (Carroll, 2014). Models demonstrate applicability in pedagogy where teachers may use them at school to diversify the learning/teaching process to motivate students. Students learn from their mistakes and through failure, realize that there are not always right or wrong solutions to many problems. The students learn to listen to each other's ideas and opinions, incorporating them into a more holistic understanding than they might otherwise develop on their own (Luka, 2020).

Design Thinking in STEM Education: Recent developments in STEM education indicate that design thinking may be utilized to augment traditional curricula, where students use the problem-solving process in the execution of a STEM project-based activity. Li, et al (2019, p. 100), argue that, while the fields of engineering and technology introduced the application of design thinking as a problem-solving tool, it should also be used in teaching and learning mathematics. Some anecdotal evidence indicates that design thinking applied to other disciplines has a positive impact on learning (Pande & Bharathi, 2020). The Diamond Afterschool Project paired graduate engineering students with middle school students as mentors for a 10-week after school program. Through informal programs, designed in collaboration between the middle-school students and their collegiate mentors, the middle school students seemed to more directly connect with the STEM topics the group explored (Carroll, 2014). Henriksen (2017) proposes that through design thinking, educators may be able to fully accomplish Science, Technology, Engineering, Arts, and Mathematics (STEAM), where the “A” goes beyond arts integration. Citing a Spanish teacher who developed a project analyzing clean water issues, problematic in Spanish-speaking countries, this teacher used the Stanford model to facilitate her students understanding the water cycle problem and illustrating their solution graphically, tying earth science and art to language (Henriksen, 2017). Further, results of existing studies focused on engineering and technology indicate promise for design thinking, systematic studies looking at more STEM fields would provide important understanding at its broader application (Li et al., 2019).

Implementing Design Thinking: The Learning Design Studio, based upon the design inquiry model, combines the process of design thinking with the constructivist theories of inquiry learning. It creates an environment where students and teachers employ a structure, which upon examination appears linear but in reality, the project work is messy and iterative (Mor & Mogilevsky, 2013). The quality of the design process and the resulting solution is dependent on the expertise of the teacher or facilitator. The implementation of the project must be carefully developed to suit not only the ability of the students, but also the skill of the facilitator. (Mosley et al., 2018). Several challenges appeared in the Swinburne University course the first time it was taught (Melles et al., 2012). One mentioned by the authors that seems to be consistent is that there is inadequate time during the project for the students to fully develop and test their design proposals (Melles et al., 2012). Mor and Mogilevsky (2013, p. 14) note that teachers "find it difficult to bind these to pedagogical theory. To counter this tendency and to allow them to retain their tacit pragmatic knowledge while adopting an attitude of scientific rigor, they need to acknowledge their role as designers of learning experiences". The implementation of design thinking requires a fundamental paradigm shift for educators who must be specifically trained in design thinking methodology, create projects that allow students the time and directed inquiry that fosters the use of the technique (Leverenz, 2014).

Theoretical Frameworks

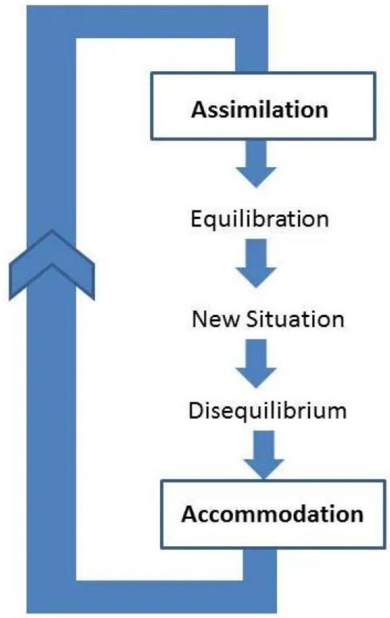
Constructivism in Education

The theory of constructivism in education has developed several branches and numerous definitions in the decades since it was first proposed by Piaget early in the twentieth century. Saul McLeod (2018, p. 4) summarized Piaget's precepts that "cognitive development was a progressive reorganization of mental processes as a result of biological maturation and environmental experience". Piaget further proposed that children gain experience in the world and build new knowledge through observations upon previous knowledge (McLeod, 2018, June 06). This concept became the foundation for the "spiral curriculum" first proposed by Jerome Bruner (1960). This branch is known as "cognitive constructivism" where cognitive development happens because the child builds or constructs their own knowledge through independent exploration (McLeod, 2018, June 06). This theory supposes that the learner builds new knowledge on top of previous knowledge, e.g: by applying knowledge of addition and subtraction previously acquired, the student learns to solve multiplication and division.

The iterative nature of design thinking parallels the Constructivist principle of applying a spiral curriculum (Bruner, 1961) to learning where new knowledge is built upon a foundation of knowledge previously acquired. Each iteration adds knowledge and informs a more refined choice of guidelines by which the designer is ideating possible solutions. Designers construct their knowledge of a problem and potential solution with each iteration. These subsequent iterations and built knowledge then form the experience used by the designer to adjust their guidelines. The basis of this assumption lies in Jean Piaget's assertion of the Process of Adaptation (Piaget & Cook, 1957; McLeod, June 2018), where a student assimilates new understanding into their existing body of knowledge. New knowledge is accommodated into their existing body of knowledge and existing understanding is adjusted (equilibrated) to incorporate this new knowledge, adjusting and balancing all their knowledge (Piaget & Cook, 1957). They are then introduced to additional new knowledge and the process iterates (McLeod, June 2018). (see Fig. 2-3)

Figure 2.3

Piaget's Process of Adaptation



Note: Diagram illustrating Piaget's Process of Adaptation (McLeod, 2018, June 06)

Jerome Bruner's constructivist theory built upon Piaget, proposing learning is an active and self-initiated process. The learner 'constructs' new knowledge on top of existing knowledge. New information is transformed and the learner develops hypotheses to make decisions (Bruner, 1960, p. 33) Further, the learner continues to build a mental model allowing them to project or synthesize beyond the information observed (Culata, 2018) reinforcing the parallels with design thinking. Observations made in the testing phase add new information. This is assimilated into the designer's collective experience (total knowledge) and a revised prototype is ideated (revised hypotheses) which is further tested and observed, reconstructing a more revised cognitive structure.

Historically, design thinking education has been focused in secondary and graduate education settings on the premise that a large base of knowledge is needed to frame the problem and establish guidelines for ideating a proposed solution (Pande & Bharathi, 2020). Bruner posed the "hypothesis that any subject can be taught effectively in some intellectually honest form to any child at any stage of development" (Bruner, 1960, p. 33). This statement may be interpreted to mean that all subjects can be introduced to students of any age. The concepts and terminology may have to be simplified or taught in a very fundamental introductory way that is appropriate to the age. Bruner contends that there are foundations of all subjects that can be laid, even at an early age, before the child is developmentally ready for the complex theoretical underpinnings to the concept. Accepting this hypothesis, it is conceivable that a student at any age has enough experience and base of knowledge to tackle a piece of any developmentally appropriate problem.

The historically accepted premise in mathematics is that students learn procedures that have already been established in mathematics disciplines. The development or discovery of new concepts and procedures in mathematics is the realm of trained mathematicians. By Bruner's hypothesis, it is conceivable that any student with some foundation in mathematics could develop a new procedure or understanding that falls outside of those taught in institutionalized mathematics instruction (Bruner, 1960). This is reinforced by recognizing that, as the discipline of mathematics evolved from ancient times through the scientific revolution, these new disciplines were built using this process. For example, Isaac Newton developed calculus out of the previous knowledge that he had compiled with new observations.

In contrast to Piaget, Lev Vygotsky felt that student learning was a result of the cultural context in which the learner was positioned, not defined by it (Vygotsky, 1978).

A second branch of constructivism, “social constructivism,” has a foundation laid by Vtgotsky, built upon Piaget’s work where the child’s learning happens as a result of interaction with others in a social setting (Vygotsky, 1978). Vygotsky disagrees with Piaget in that learning does not happen through independent exploration, but must happen in a social context (McLeod, 2018, August 05). “Learning awakens a variety of internal developmental processes that are able to operate only when the child is interacting with people in his environment and in cooperation with his peers. Once these processes are internalized, they become part of the child's independent developmental achievement” (Vtgotsky, 1978, p. 90). This theory becomes the foundation of small-group active learning where each student brings unique understanding to the group. By sharing and collaborating their perspectives, all students acquire a deeper understanding of the material. Like Bruner, Vygotsky proposed that active learning most successfully occurred when the learner was immersed in a social context, where environment was key to learning (McLeod, August 2018). Vygotsky's theories also support the successes seen in collaborative learning. Learners placed into a group with others who have different and perhaps more advanced knowledge and led by a teacher who facilitates interaction between the students on the subject will grow their knowledge as a result (Vygotsky, 1978).

Expanding on Vygotsky’s views on the cultural context of learning, Bruner proposed that education formed the learner’s culture. He felt that this was not a one-way relationship; while the culture defined learning, the culture evolved as learning took place. Thereby, the student grows and changes as their knowledge expands, contributing to the social environment in which they are learning. With this theory, as a student learns more about topics outside of their culture, the more that they become absorbed into that culture (Bruner, 1996), that is, the more that a student learns about science, the more the student becomes a scientist.

By merging pieces of both Piaget's cognitive constructivism and Vygotsky's social constructivism, Jerome Bruner and other education theorists built the foundation upon which much of the active learning and project-based learning methodologies in use today are based. Bruner's spiral curriculum utilizes Piaget's notion that all new knowledge is layered upon previous experience which forms the framework for learning. "A curriculum as it develops should revisit these basic ideas repeatedly, building upon them until the student has grasped the full formal apparatus that goes with them" (Bruner, 1960, p. 13). For example, when teaching the concept of problem-solving with two-step equations, students are reminded of how to solve a one-step equation in that they are undoing the given operation of addition or subtraction and multiplication or division.

According to Bruner, education should always loop back, adding to existing knowledge from a fresh perspective while being structured so that it can be easily understood and processed by the student (Culatta, 2018). Building upon this philosophy, Bruner (1960, p. 33) proposed that "any subject can be taught effectively in some intellectually honest form to any child at any stage of development." This statement may be interpreted to mean that all subjects can be introduced to students of any age. The concepts and terminology may have to be simplified or taught in a very fundamental introductory way that is appropriate to the age, however, Bruner contends that there are foundations of all subjects that can be laid, even at an early age, before the child is developmentally ready for the complex theoretical underpinnings to the concept.

Further, Bruner grew Vygotsky's social constructivism when he proposed that learning is culture and happens within a cultural setting. Through education, we indoctrinate learners into a culture and that true education does not happen outside of the context of their culture or the social and family environment in which they are situated. That is, the idea that using a student's culture as a framework upon which to build the learning offers the possibility of connecting a student to topics that may historically fall outside of that student's culture. By using cultural tools and references to make the topic relevant, they can be brought to understand the principles of the topic (Bruner, 1996). A project-based learning activity is an example of a strategy built on this theory. In a PBL, a real-world problem relevant to the student is solved using new ideas or concepts and the student masters those concepts through the project.

The constructivist principle of a student constructing their understanding based upon personal experience is particularly effective in teaching visual arts. When studying the arts, students acquire knowledge and skills through a guided process of free exploration where students' content evolves through activity, thereby constructing new knowledge (Tomljenović & Tatalović, 2020, p. 28).

Discovery Learning Theory

In 1961, Jerome Bruner published an article in *Harvard Educational Review* introducing his idea of learning through discovery (Bruner, 1961), defining discovery as "a matter of rearranging or transforming evidence in such a way that one is enabled to go beyond the evidence so assembled to new insights" (Bruner, 1961, p. 2). Bruner's later theory of Discovery Learning developed as an extension of constructivism and became the practical application employed in guided learning inquiry. With discovery learning, a learner can be presented a problem or an end result with no clear connection or explanation. The learner discovers the nature of the problem as a result of observation and experimentation (Bruner, 1961) This is a parallel premise in which design thinking is applied. Emphasizing discovery in learning allows the student to construct their own learning, constructing knowledge as they organize found material. Because the learner engages in the act of discovery, they are not learning about a topic and are punished if they do not comprehend it. Instead, the reward becomes the satisfaction of discovering (Bruner, 1961, p. 5). The process of discovery learning begins with a dialogue between student and teacher where the student internalizes a base of knowledge and “rules of generation” allowing the student to begin discovering new knowledge and layering it on to their existing knowledge. In this model, Bruner concludes that the student is in a position to “experience successes and failures, not as reward and punishment, but as information to be processed and added to the base of knowledge” (Bruner, 1961, p. 6). Finally, Bruner proposes that through the act of discovery, the knowledge is firmly planted in memory.

Since the introduction of the theory by Bruner in 1961, discovery learning has come under criticism, questioning the effectiveness of the concept. If the discovery learning is intended to occur when students are presented with material and data with no direction of clear objective, this is considered unassisted discovery learning. Historically, criticism of discovery learning has mostly grown from analysis of this technique. When compared to unassisted discovery learning, explicit instruction is consistently shown to be more effective (Alfieri et al., 2011, 11). However, implementing unassisted discovery contradicts the premise proposed by Bruner where the student is given a structure and guidance in which to explore (Bruner 1961, p. 6).

An implementation of discovery which more closely aligns with Bruner's proposed structure (Bruner 1961, p. 6) is enhanced discovery learning. With enhanced discovery, a teacher or facilitator provides an explanation, learners are trained on exploration techniques, and then the teacher monitors and prompts student progress through the discovery process. Studies suggest that with this model, learners take more time to synthesize material learned and subsequently scored higher on post-tests. (Alfieri et al., 2011, p. 11).

A key advantage noted in the implementation of discovery learning is the ability of students to visualize a concept. After a project using origami to teach geometry, students commented that they could more readily see the object because of using the method and thought that it was easier to complete questions on their worksheet because of an intimate knowledge of triangles. Another opportunity in discovery learning lies in the implementation of group activities. Teachers observed increased student engagement because of the collaboration with classmates (Maarif, S. 2016, p. 119). Some evidence suggests that discovery learning techniques benefit older learners more than adolescents who require more thorough explanation or explicit instruction before undertaking exploration activities (Alfieri et al., 2011). While unassisted discovery has shown to be less effective, when guided through the inquiry process, students engaged in discovery learning have experienced notable success (Alfieri et al., 2011). Discovery learning and collaborative learning have become the theoretical foundations upon which project-based learning has been developed. Based on this theoretical framework, it is proposed that within the pedagogy of guided inquiry project-based activity, students can solve a complex, multi-variable problem using design thinking, and based upon their collective existing base of knowledge expand their knowledge about a given topic.

Challenges of Implementing Constructivist and Discovery Learning Theories

Raising curiosity in the students and equipping them with the skills and confidence needed to undertake independent learning is a prerequisite to implementing a discovery activity (Bruner, 1961). According to Jerome Bruner, discovery is rarely the frontier of knowledge, but instead, "in its essence a matter of rearranging or transforming evidence in such a way that one is enabled to go beyond the evidence" (Bruner, 1961, p. 2). Therefore, to implement discovery learning, the student must be taught to identify evidence and arrange it into a new and meaningful pattern.

Since evidence suggests limited efficacy from unstructured and unassisted discovery learning activities, substantial preparation is required of the teacher to assure that the lesson requires "learners to be actively engaged and constructive" (Alfieri et al., 2011, p. 13). To be effective, tasks must be carefully scaffolded to guide students through the process, using worked examples of how to successfully complete the task. Students must learn to explain their own ideas and learn to test those ideas against the evidence. Finally, students must be taught that failure is a useful and unavoidable part of the discovery learning process (Alfieri et al., 2011, p. 13). Another significant hurdle may be that the discovery process takes considerably more time to work through than explicit instruction. Students must be given ample time to make observations and draw conclusions, testing their ideas against the evidence and then documenting their findings (Tomljenović & Tatalović, 2020, p. 28).

To teach these skills, the instructor must become fluent in the concepts and techniques of discovery. This requires teachers to retool the teaching techniques and education paradigms that most of them were taught in college. According to Tomljenović & Tatalović (2020, p. 29), this is the single biggest reason that teaching is still often reduced to explicit instruction sessions “instead of more complex ones through student collaboration. Student-initiated questions and student-to-student interactions are neglected, conventional knowledge and ways of thinking are preferred, or the curriculum is interpreted in a rigid and inflexible manner” (Tomljenović & Tatalović, 2020, p. 29).

Summary

Many adolescent African American students in underserved communities lack success in STEM content. Some reasons for this include their disinterest and apathy toward the subject matter, being underprepared as they enter middle school, as well as having unqualified teachers instructing them in mathematics and science specifically. African American students can experience improved performance in STEM by engaging in collaborative arts integrated, design thinking, project-based learning experiences. These constructs provide structure, movement, and out-of-the-box thinking for this student demographic to explore. Constructivist and Discovery Learning methodologies provide a foundational blueprint for instructors to follow that allow students to fully engage in culturally relevant lessons.

Chapter 3
Methodology

Introduction

Chapter 3 outlines the methods that were utilized to answer the research questions and test hypotheses surrounding the efficacy of project-based learning (PBL), solved with design-thinking and expressed in art or DAIP, to teach principles of mathematics. While some anecdotal evidence suggests that PBL could help to improve student engagement and understanding, (Carroll, 2014), little quantified evidence exists to substantiate the use of design thinking in middle school STEM classrooms (Pande & Bharathi, 2020). The research methodology was designed to quantify the effectiveness of an alternative instructional approach in improving academic performance and increasing interest and confidence in African American students who otherwise may feel disenfranchised from success when studying pre-algebra.

Research Questions

1. To what extent would minority students' STEM scores in a disenfranchised community improve when taught STEM concepts using an alternative problem-solving teaching methodology (DAIP)?
2. To what degree would minority students in a disenfranchised community improve in confidence and interest in STEM concepts when taught mathematics using an alternative problem-solving strategy (DAIP)?

Hypotheses

$H_0 1$: There is no significant difference in the achievement in pre-algebra as measured by the Renaissance STAR Benchmark Assessment (Renaissance, 1998) and a teacher-generated pre- and post-assessment between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

$H_a 1$: There is a significant difference in the achievement in pre-algebra as measured by the Renaissance STAR Benchmark Assessment (Renaissance, 1998) and a teacher-generated pre- and post-assessment between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

$H_0 2$: There is no significant difference between the confidence and interest in STEM topics as measured by scores on the Attitudes Toward Mathematics Inventory (ATMI) (Tapia, 1996) between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

H_a2: There is a significant difference between the confidence and interest in STEM topics as measured by scores on the Attitudes Toward Mathematics Inventory (ATMI) (Tapia, 1996) between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

Research Design

According to Creswell (2014), quantitative research is a post-positivist worldview where phenomena are observed and numerical data collected and analyzed to measure changes in variables. The nature of quantitative research is establishing a correlation, “association (or relationship) between two or more variables” (Creswell, 2014, p.12). As the independent variable is altered, a corresponding change in the dependent variable is measured. For example, a researcher attempts to determine a relationship between achievement and attitude by analyzing the changes in variables within a sample population. The independent variable was defined as the teaching strategy and the dependent variables were identified as student outcomes and student confidence in their mathematical ability. The variables were measured with a series of standardized tests as well as changes in student interest and confidence in the subject as measured by the Attitudes Towards Mathematics Inventory survey (Tapia, 1996).

The sample pool for the quasi-experimental design consisted of eight sections of a pre-algebra course. The student body was composed of 99% African American and 1% Other eighth-grade students at a Midwest urban middle school. A quasi-experimental design was chosen because the students were to be divided by course section, into two groups, one group receiving the intervention and the other acting as the control (Creswell, 2014). The intervention group were taught the same content using design-thinking to solve a grade-appropriate, complex problem (Leverenz, 2014). The project was developed around the students' participation in a Math-A-Thon; a problem that was relevant to the students since the entire eighth grade participated in the event. The intervention group executed a project-based learning (PBL) activity where they designed a fund-raising model and analyzed their earning potential for the Math-A-Thon. The students were allowed to express their understanding autonomously through various artforms. This PBL incorporated multiple disciplines from STEM, but centered the learning on specific mathematics skills surrounding functions as the measure. The students executing the PBL expressed their understanding of functions through artforms that they determined best represented the principles that they learned. The non-intervention group was taught the same mathematical concepts about functions using traditional empirical lessons with direct instruction and no connection to the Math-A-Thon. (Table 3.1)

Table 3.1

Research Design

Dates	Treatment Group	Control Group
9/22	Fall Renaissance STAR Benchmark	Fall Renaissance STAR Benchmark
10/22	Pre-ATMI Survey	Pre-ATMI Survey
10/22	Teacher-Generated Pre-Test	Teacher-Generated Pre-Test
10-11	**DAIP - Math-A-Thon**	**Direct instruction - Functions Unit**
12/22	Post-ATMI Survey	Post-ATMI Survey
12/22	Teacher-Generated Post-Test	Teacher-Generated Post-Test
1/23	Winter Renaissance STAR Benchmark	Winter Renaissance STAR Benchmark

Since the sample selection of the type of instruction that the students would receive was dictated by the school during enrollment, possibly the eight control and intervention groups could vary in mathematics skills and understanding. All students received the same pre- and post- assessments to control for variation in background understanding and skills. Four sections received the intervention teaching strategy and four sections received the traditional instructional methodology (Creswell, 2014). The Renaissance STAR Math assessment (Renaissance, 1998) was administered to all participants in both groups early in the school year to determine a baseline of student knowledge of the concepts to be taught through the intervention and to establish that no statistical difference with respect to background in algebraic concepts initially existed between the control and experimental. The teacher in the classes delivering the intervention was prepared in exercising the design thinking process by direct involvement in the planning of the PBL. The intervention classes were given the complex, multi-discipline problem surrounding the Math-A-Thon to solve. The non-intervention group was provided direct instruction from the teacher in the manner and style that the teacher has used for the material in previous years.

After completing the unit lessons, all eight sections took the Renaissance STAR Math assessment (Renaissance, 1998) and the teacher-generated unit test. To measure improvement in achievement, the intervention and control classes' scores were compared on the Renaissance STAR Math assessment (Renaissance, 1998) and the teacher-generated unit tests. Data analysis looked for evidence of a significant statistical difference between those who received direct instruction and those who were taught with the project-based learning strategy.

Data about student confidence and interest in STEM was gathered using the Attitudes Toward Mathematics Inventory (ATMI) survey (Tapia, 1996) using Google Forms before the assessments were administered. A series of questions were asked and students scored their responses using a five-point Likert scale. Responses between the intervention group and non-intervention group compared responses that students felt more positive and confident after being instructed by one strategy or the other.

Threats to Validity

According to Creswell (2014), maturation, selection, and mortality (removed from the study) are all internal threats to validity involving the participants. To mitigate the threat posed by maturation, all students in the sample were taken from the same middle school class at a single school. All students in the sample were approximately thirteen years old and had nearly identical experience in mathematics coursework when entering the eighth grade. Because sample pools were not random or self-selected, in September, all subjects were given the Renaissance STAR Math assessment (Renaissance, 1998) as a pre-test to establish consistent skill levels and to confirm that variation between individuals was evenly dispersed between the intervention and control groups. The Renaissance STAR Math assessment (Renaissance, 1998) is a standard exam chosen by the school district to evaluate student progress in mathematics and reading. It was used as a subject matter assessment instrument because it is standard district practice to be administered as a part of the district-wide curriculum. Using standardized test scores as a measure reduced bias which might have been injected by the teacher. A standardized test was administered to all eighth-grade students mitigating any possible threat to validity posed by the instrumentation used to measure student progress. Scores on an additional end-of-unit assessment created by the teacher and administered as part of the unit validated data collected from the Renaissance STAR Math assessment (Renaissance, 1998).

The internal factor that posed the greatest threat to validity was mortality. Because the sample size was moderate, students whose parents chose to not have their child participate or who removed their child from the study posed the potential to significantly influence the results. The greatest external factor that threatened the validity of the research was acquiring a large enough sample size, therefore, multiple mathematics sections needed to be included in the study. Student assessments were graded and survey data was collected by the regular classroom teacher.

Participants and Sampling

The participants were African American adolescents in an urban, underserved community with limited access to instructional resources. The participant pool consisted of eight sections of eighth-grade pre-algebra students at a public middle school in a Midwest urban district during the fall semester of the 2022-2023 academic year. The district's overall enrollment was approximately 5,700 students with about 790 students in the middle school. The middle school consisted of students at the sixth, seventh, and eighth-grade levels with the eighth-grade population nearing 230 students. The district demographic distribution is indicated in Table 3.2.

Table 3.2

School Demographics

Ethnic Group	Percentage of School Population
African American	95%
Asian American	2%
Caucasian	2%
Other	1%

Note: Since 2016, 100% of the students in the district receive free and reduced breakfast and lunch (Missouri Department of Elementary and Secondary Education, 2019).

A portion of the school population consists of students in transition. These individuals are homeless and either living in a shelter or with family members who reside in communities outside of district boundaries. A district-appointed *Students in Transition* (SIT) coordinator ensures that these students have transportation to and from school, access to mental health services, clothing, and uniforms, as well as hygiene products and school supplies. In addition to serving an identified population of students in transition, the school also experiences a high transiency rate which, in turn, makes it difficult for students to gain and retain information. Student performance on district, local, and state assessments are indicative of the lack of educational consistency many students face.

Student mathematics placement is based on local, district, and state assessment scores and academic performance in mathematics in prior years, specifically at the fifth, sixth, and seventh grade levels. Students who excel in mathematics are placed in Algebra I at the eighth-grade level and those who pass with a B or better qualify to take the state

End of Course Assessment (EOC) and receive high school credit at the end of their eighth-grade year. Students who do not qualify for Algebra I are placed in pre-algebra.

The school organizes eighth-grade students in two teams: Team A and Team B. Students are assigned to a team randomly upon enrollment with the intention of keeping the numbers balanced between teams of teachers. Under ideal circumstances, each team has one mathematics teacher, however, due to a shortage of mathematics teachers, both teams were taught by the same teacher concurrently when the data was collected. Each section met for 57 minutes per day. The math teacher created pre- and post-assessments for all units of instruction as outlined in the district curriculum.

The pre-algebra teacher instructed multiple sections of students. The instructor was a 21-year veteran of the school who had not had prior experience with design thinking strategies or its implementation. This teacher regularly attended professional development provided by the school and sought opportunities outside the district to employ new ways of challenging students' critical thinking skills, interest levels, and understanding of difficult mathematical concepts. Students in the treatment group engaged in a project-based learning activity integrating arts and discovery learning and students in the control group followed a traditional method of instruction for one unit of study over the course of eight weeks of instruction. (Table 3.3)

Table 3.3

Research Population Design for Comparison of PBL to Traditional Instruction

Groups	Number of Sections	Number of Students per Section	Number of Students in the Sample
Control	4	9	36
Treatment	4	9	36

Note: Table represents the sample distribution between two groups, N = 72.

Creation of the Design Arts Integrated Project

The development of the Design Arts Integrated Project (DAIP) activity was key to the success of this research. The DAIP had to engage the students while meeting all necessary district and state standards knowing the responsibility of meeting these standards primarily fell to the classroom teacher. The unit of study explored consisted of eight weeks covering functions. Since project-based learning, the foundation of DAIP, is a strategy that was new to this particular classroom teacher, the researchers collaborated with the teacher to create an activity that retained creativity, employed design thinking to solve the problems, guided the artistic expression of the result by the students, and provided a means of assessing student progress and learning.

The DAIP activity was developed during the summer of 2022 outside of the teachers' normal preparation blocks. Through the process of collaborating with the researchers on the creation of the activity, the classroom teacher became intimately familiar with implementing project-based learning, design thinking, and identifying possible artistic expressions that could be used as they coached the students when the intervention was implemented in Fall 2022. With minimal interactions from the researchers once the unit began, the intention of this model was to provide the most authentic experience for both the students and teacher in deploying the strategy.

Instrumentation

The Renaissance STAR Math Assessment (Renaissance, 1998) is a "scaled, integrated learning platform used by approximately 16.7 million students in approximately 45,000 schools worldwide" (Wyse et al., 2020, p. 1). STAR assessments were constructed to link to standards in every state and the results indicate national norms against which individual student performances are compared (Wyse et al., 2020, p. 2). The National Center on Intensive Intervention (2018) funded by the U.S. Department of Education, confirmed that the Renaissance STAR Math Assessment (Renaissance, 1998) aligned to state-specific standards and the Common Core State Standards. The assessment was designed to test four major areas of student understanding: Numbers and Operations, Algebra, Geometry and Measurements, and Data Analysis Statistics and Probability (National Center for Intensive Intervention, 2018). To confirm a correlation of student performance with state standards, a large sample (n = 131,103) of students completed STAR Math assessments and returned a median coefficient of reliability of 0.93. To verify the validity, a sample size of 4000 eighth-grade students took the test concurrently with their eighth- grade mathematics course and their scores returned correlation of 0.75 compared to state standards (National Center for Intensive Intervention, 2018).

The Attitudes Toward Mathematics Inventory (ATMI) developed by Martha Tapia from Berry College (Tapia, 1996) consists of 40 questions designed to measure attitudes toward mathematics in high school and college students. Designed to be administered to middle and high school students this instrument was developed to investigate students' attitudes toward mathematics. According to the author, unlike other surveys on attitudes toward mathematics, the ATMI was "designed to be brief while also capturing multiple factors that contribute to one's attitude about math" (Tapia, 1996, p. 11). To validate reliability, Cronbach's alpha was calculated to be 0.963. After the author deleted questions 13, 36, 39, 40, 43, 44, 45, 46, and 47 alpha increased to a value of 0.9667 (Tapia, 1996, p. 11). Content validity was established by having the questions reviewed by impartial experienced teachers. Construct validity was confirmed by demonstrating that all questions had an item-to-total correlation higher than 0.49. "Another method that was used to determine construct validity was by using factor analysis" (Tapia, 1996, p. 11-12). These results indicated that the instrument is measuring only one construct.

Data Collection Procedures

For every unit of study in the district math curriculum, a specific set of state standards is taught and a collaborative teacher-generated pre- and post-test is administered to students to measure academic growth on a specific unit covered over a predetermined number of weeks. Each question response was charted for the conduction of error analysis resulting in students being provided explicit instruction in identified high-needs areas. The Attitudes Towards Mathematics Inventory (ATMI) survey was administered to all sample groups prior to the pre-test and again prior to taking the post-test. The survey was given prior to participants taking both the pre- and post teacher-generated assessments to mitigate negative responses due to the stress of test-taking.

The Renaissance STAR Math assessment (Renaissance, 1998) is utilized as a district-wide benchmark assessment given three times per year: fall, winter, and spring. This grade-level assessment measures proficiency on every math standard defined by the state in which it is administered. All eighth-grade students are required to take this assessment regardless of math placement. Students were administered the Renaissance STAR Math assessment by their classroom teachers in September and again in January (Renaissance, 1998).

The data from the September test provided a baseline of student knowledge before implementing the intervention unit of study. Relevant sections of the January test were compared to the September test to determine growth in student knowledge from the intervention. The results of the impact of the design thinking, arts-integrated intervention implemented with the treatment group was evaluated in each assessment.

The intervention, a project-based learning activity (Appendix A), was developed by the regular classroom teacher under the direction of the researchers to assure that all curricular standards were met by the lesson. The regular classroom teacher delivered all instructional materials, developed and administered the pre- and post-tests, proctored surveys, and administered the Renaissance STAR Math exam in conjunction with the school district standard policy. The teacher provided the researchers with paired results, including test scores and survey responses.

Data Analysis Procedures

The scores on the September Renaissance STAR Math exam and the teacher-generated test were compared using a two-sample, two-tailed t-test to determine that no statistically significant difference exists between the groups prior to the intervention. Likewise, the responses for both groups on the ATMI pre-survey were compared using a two-sample two-tailed t-test to determine that no statistically significant difference exists between the groups before the intervention. A two-sample, two-tailed t-test was chosen because it allowed the Treatment and Control groups to be compared in both directions, above and below the mean.

After the completion of the eight-week activity, each subject’s September and January Renaissance STAR Math exam scores and pre- and post teacher-generated test scores were analyzed for change using paired t-tests. Likewise, each sample group’s pre-survey and post-survey responses were analyzed for change using a paired t-test.

All test scores on the pre-test for the Treatment group and all scores for the Control group pre-test scores were then merged into their respective pools. An Analysis of Covariance (ANCOVA) was used to determine whether there were any significant differences between the Control and Treatment groups on the dependent variables: teacher-generated test, ATMI survey, Renaissance STAR Math exam (Laerd, 2018). An ANCOVA was chosen because it assesses the extent to which an independent, categorical variable (teaching strategy or DAIP) is associated with statistically significant differences in a continuous, dependent variable (student outcomes and student confidence) while controlling for a third variable called the covariate (baseline achievement in mathematics as measured by pre-assessment scores on Renaissance STAR and teacher-generated tests) in order to remove the effect of the covariate on the relationship between the independent and dependent variables (Laerd, 2018). When the sample size is large, normality in the data distribution is less critical to achieve adequate power in parametric tests such as an ANCOVA or t-test. When the sample size is small, parametric tests do not have adequate power to detect small differences and require normally distributed data. If the data, or the residuals in the case of ANCOVA, are not normally distributed, a non-parametric test is required to verify the results of the parametric test (Cody, 2021. pp. 88-89).

Since the ATMI survey consists of forty questions with responses on a 5-point Likert scale and these responses represent ordinal data, parametric analysis is not appropriate on individual questions. Instead, the questions were grouped by topic or theme and individual responses to the questions in a group averaged. This average could then be more meaningfully analyzed with a t-test or ANCOVA (Joshi et al., 2015. pp. 399-400). Questions were also analyzed using a non-parametric test to accommodate the ordinal data.

Generalized shifts in student learning, interest, and confidence were identified correlating to the efficacy of the teaching strategy. If a subject withdrew from the study prior to the implementation of both the pre- and the post- administration of any one instrument, that subject's scores were withdrawn from the sample.

Ethical Considerations

All participants involved in the research were eighth-grade students. A strict adherence to the guidelines approved by the Institutional Review Board were followed. In advance of the study, parent/guardian consent and student assent forms were developed, explained, distributed, and collected from each willing participant by the classroom teacher and researchers. Students were explicitly informed that their involvement was strictly voluntary and that they would receive grades based solely upon their academic performance on the functions unit without influence from participating in the research. Further, it was explained that all eighth-grade students would be learning the same concepts about functions. Functions are a regular topic taught in eighth grade and would not be a topic taught out of the usual curriculum timeline.

The teacher, a 21-year veteran mathematics teacher who had been working at this grade level in the school for several years, was briefed on the objective of the study and provided design-thinking professional development. The importance of maintaining participant confidentiality and the non-disclosure of Treatment and Control group members was stressed to the teacher that any divulgence could be a threat to validity. All assessment and survey results were gathered electronically, thereby minimizing the risk that the data could be compromised and viewed by unintended parties.

Each student was assigned an identification code by the researchers to maintain anonymity; allowing the researchers to track individual responses for accurate data collection and analysis. As previously stated, all assessments and surveys were conducted online. Measures were put in place to ensure data was untraceable, therefore, participants were only permitted to use school-issued devices on a secure campus network when completing teacher-generated pre- and post-tests, Renaissance STAR Math assessment (Renaissance, 1998) and interest/confidence surveys. The collection and analysis of this data did not present a significant risk to the physical, mental, social, emotional, legal, or overall well-being of the participants.

Limitations

The class enrollments for each section were predetermined by the school at the beginning of the school year. One teacher was responsible for instructing all sections of pre-algebra at the eighth-grade level and was chosen to participate as the classroom teacher. Prior to implementing the alternative teaching strategy, this teacher had no experience with project-based learning or design thinking. The researchers provided instruction to the teacher and collaborated with that teacher in the development of the DAIP activity to ensure that they were prepared to effectively deliver the lesson.

Of a potential pool of approximately 200 pre-algebra students enrolled at the eighth grade level, 72 students and their guardians signed assent and consent forms before the data was collected. While all eight sections of pre-algebra participated in either the Treatment or Control group activities, the data was parsed and only the results for the 72 in the study were analyzed. The modest final sample size likely factored into the inconclusive findings. In addition, a small percentage of the sample pool were eligible for special education services. Because these students would have been working in collaboration with special education teachers, these teachers may have altered the students' performance due to accommodations that were made in the project.

Delimitations

The research was conducted in a single grade within a single middle school in an urban district. Since the research targeted African American students, the school chosen had a population that was nearly100-percent African American, meeting the demographic requirements for the research. A quasi-experimental design was implemented because samples were not selected at random or self-selected. Instead, the source sections were assigned by the school during the enrollment process. The researchers chose to apply the intervention to a single eight-week unit of study covering functions in the fall semester of the pre-algebra course. The pre-algebra courses had the highest enrollment and gave the largest potential pool of subjects. All students in the eight sections of pre-algebra were solicited to participate, with 72 agreeing and signing assent and consent forms. The samples were chosen for convenience based on student enrollment in eight sections and all sections were taught by a single teacher. The researchers relied on the regular classroom teacher of the pre-algebra sections to deliver all instruction and assessments, thereby removing themselves from direct involvement in the intervention.

Summary

A direct comparison between teaching pre-algebra using direct, explicit instruction and an alternative teaching strategy, DAIP where the students employ design thinking to solve a complex problem has the potential of providing teachers of marginalized students alternative means of engaging them in STEM subjects. Little quantitative evidence previously existed as to the effectiveness of this strategy (Pande & Bharathi, 2020), but by quantifying the results of both student achievement and student interest in mathematics, teachers may be convinced to invest the time and effort in implementing such an alternative strategy.

To answer the research questions, the sample pool of 72 was divided into two groups; Treatment and Control. A series of two instruments was used to explore the first question; the Renaissance STAR standardized exam and a teacher-generated assessment. To address the second research question, the Attitudes Towards Mathematics Inventory (ATMI), a forty-question survey answered on a five-point Likert scale was used. Each instrument was given prior to the start of the functions unit of study and again at the completion of the unit.

Scores from all instruments administered prior to the intervention and following the intervention were collected and comprised the data. Eleven competencies were parsed from the standardized exam that specifically pertained to functions. Scores from each of these competencies were averaged to create a composite score for each student on the September exam and a separate composite score on the January exam. Raw scores from the teacher-generated assessment were compared without modification. The forty questions on the ATMI survey address four key subscales that pertain to attitudes and confidence in learning mathematics. Responses to the questions were compiled into the four subscales and an average, composite score was generated for each respondent for each subscale.

In the first phase, the pre-test scores were used to determine a baseline achievement level in pre-algebra, specifically on functions to establish whether the Treatment and Control groups had similar achievement in functions prior to the intervention. Scores on the pre-test for the Treatment and Control groups were compared using an independent-samples t-test.

In the second phase, composite scores from the Treatment and Control groups' pre- and post- scores for each instrument were then compared. To determine if a statistically significant difference existed between the groups, scores from the pre-test and post-test were compared using an Analysis of Covariance (ANCOVA) between the Treatment and Control groups. If the data, or the residuals in the case of ANCOVA, were not normally distributed, a non-parametric test was required to verify the results of the parametric test (Cody, 2021. pp. 88-89).

Chapter 4
Findings

Introduction

Opportunities for African Americans to have equal access to education have been beset with barriers since the earliest years that education opportunities existed following emancipation (Allen and Jewell, 1995). Racism and restricted freedoms experienced by African Americans have led to drastically different learning opportunities and vast inequities in student achievement (Darling-Hammond, 2000). As a result, many students in urban and suburban school systems experience a disparity in academic resources solely due to the community in which they reside (Jenkins, 2017). Consequently, some African American students struggle in classrooms where explicit or direct instruction as science, technology, engineering, and mathematics (STEM) education are traditionally taught. Disengaged from STEM topics, these students feel incapable of succeeding in STEM fields (Johnson & Kritsonis, 2006). Instead, they more clearly express knowledge through alternative, more artistic media (Jenkins, 2017). Using design thinking to solve a project-based activity and expressing their understanding using the arts may be a means of capturing and retaining African American students' interest in STEM subjects.

Project-based learning (PBL), a component of DAIP, is a classroom instructional strategy designed to engage students in fundamental principles by investigating real-world problems (Cervantes et al., 2015). While executing a PBL, learners develop skills of collaboration, creative-thinking, problem-solving, and communication while solving multi-faceted problems by creating a project (Bell, 2010; Duch et al., 2001; Hmelo-Silver, 2004; Jonassen, 2000, 2004; Savery, 2006; Sendag & Ferhan Odabasi, 2009). The DAIP activity required students to develop a plan for participating in a Math-A-Thon to raise funds for a charity devoted to helping cancer-stricken children. The specific pre-algebra competencies explored in the activity focused on the study of functions, where every fund-raising pledge was analyzed and expressed as a function. Chapter four provides the results of research measuring the effects that DAIP which incorporates project-based learning (PBL), using design thinking with learning expressed in an arts component, had on middle student achievement in pre-algebra and their resulting attitudes and confidence toward learning mathematics. Two research questions frame the findings.

Research Questions

1. To what extent would minority students' STEM scores in a disenfranchised community improve when taught STEM concepts using an alternative problem-solving teaching methodology (DAIP)?
2. To what degree would minority students in a disenfranchised community improve in confidence and interest in STEM concepts when taught mathematics using an alternative problem-solving strategy (DAIP)?

The quantitative statistical analysis of the results from pre- and post-tests as well as a pre- and post-survey addresses the null hypotheses developed for each research question by comparing the dependent variable values between treatment and control groups.

Hypotheses

H_01: There is no significant difference in the achievement in pre-algebra as measured by the Renaissance STAR Benchmark Assessment (Renaissance, 1998) and a teacher-created pre- and post-assessment between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform and those students experiencing a traditional pre-algebra curriculum.

H_a1: There is a significant difference in the achievement in pre-algebra as measured by the Renaissance STAR Benchmark Assessment (Renaissance, 1998) and a teacher-created pre- and post-assessment between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

H_02: There is no significant difference between the confidence and interest in STEM topics as measured by scores on the Attitudes Toward Mathematics Inventory (ATMI) (Tapia, 1996) between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

H_a2: There is a significant difference between the confidence and interest in STEM topics as measured by scores on the Attitudes Toward Mathematics Inventory (ATMI) (Tapia, 1996) between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

Data Description

Three instruments were employed to measure student progress: performance on a standardized assessment, the Renaissance STAR exam administered at the start of the school year and again at the end of the first semester, a teacher-generated pre-test and post-test, and responses on the Attitudes Toward Mathematics Inventory (ATMI) survey.

The data was obtained from a sample of 72 eighth-grade students enrolled in pre-algebra at a public middle school in a Midwest urban district during the fall semester of the 2022-2023 academic year. The participant pool consisted of eight sections of students. The Treatment group consisted of 36 students enrolled in the sections that performed the DAIP and the Control group consisted of 36 students enrolled in a traditional, direct-instruction pre-algebra course. Scores on a teacher-generated pre-test and post-test as well as the Renaissance STAR exam were used to measure achievement in pre-algebra on the topic of functions.

Additionally, the Attitudes Toward Mathematics Inventory (ATMI) was administered before the pre-test and the post-test and responses were used to measure student confidence and interest in learning mathematics. All data collected were entered into an Excel spreadsheet and checked for errors. Afterward, the data were imported into Statistical Analysis Software (SAS) for descriptive and inferential statistical analysis.

Data Analysis

Data was collected in two phases of the research. Before the functions unit in the pre-algebra classes, identical standardized assessments and teacher-generated pre-tests were administered to both the Treatment and Control groups. In addition, the Attitudes

Towards Mathematics Inventory (ATMI) survey was given to both the control and treatment groups.

Following the functions unit, both groups completed a second copy of the ATMI survey and received an identical teacher-generated post-test. At the end of the semester, all eighth-grade students retook the standardized Renaissance STAR assessment. Scores from all instruments administered prior to the intervention and following the intervention were collected and comprised the data. For the analysis of all data, a confidence interval of 95% and a type I error rate of .05 were used to interpret all statistical results.

Renaissance STAR Exam

Scores from the Renaissance STAR exam were collected for exams given in September and January. Eleven competencies were parsed from the standardized exam that specifically pertained to functions. Scores from each of these competencies were averaged to create a composite score for each student on the September exam and a separate composite score on the January exam.

One student from the initial sample population did not return after winter break and did not take the post-test Renaissance STAR, despite having completed all instruments before the intervention, participating in the activity, and taking the post-teacher-generated test and post-survey. This one student's scores on the Renaissance STAR were excluded, resulting in an N=71, however, their results were included in the teacher-generated assessment and survey data. In the first phase, scores from the September exam were used to determine a baseline achievement level in pre-algebra, specifically on functions. To establish whether the Treatment and Control groups had similar achievement in functions prior to the intervention, composite scores on the

September exam for the Treatment and Control groups were compared using an independent-samples t-test. This test compared the pre-algebra functions scores of students enrolled in DAIP with students enrolled in sections teaching pre-algebra using traditional direct instruction. The results are reflected in Table 4.1 and indicate that students in the Treatment group had a marginally higher score (M = 15.15, SD = 14.65) than students in the Control group (M = 13.44, SD = 10.54), p = 0.57.

Table 4.1

Renaissance STAR Pre-test Results

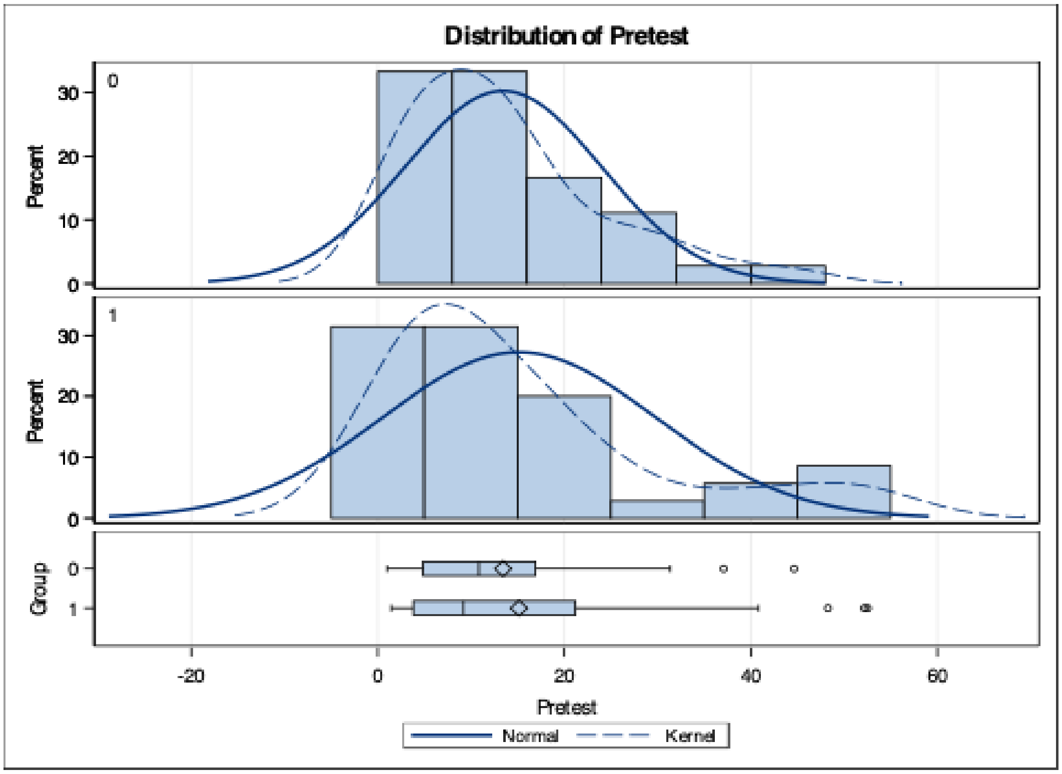
Group	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	36	13.44	10.54	1.76	1.0000	44.64
Treatment	35	15.15	14.65	2.48	1.45	52.55

Note: Pooled p = 0.57 > 0.05 Satterthwaite p = 0.58 > 0.05 (Appendix D, p.171)

The t-test results indicated that the data was not normally distributed as indicated in Figure 4.1. The box plots below the histogram indicate that the two groups significantly overlap with the mean of the Treatment group slightly higher than that of the Control group.

Figure 4.1

Data Distribution for Renaissance STAR Pretest



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. (Appendix D, p.172)

Because the data did not appear to be normally distributed, a parametric test does not have adequate power to accurately predict significance, and a non-parametric Wilcoxon Rank Sums test was also administered. (Table 4.2) The results for the Wilcoxon Rank Sums test reflected in Table 4.2 indicate that the Sum of Scores for the Treatment group (SoS=1250.0, Expected = 1260.0) is similar to the Control group (SoS=1306.0, Expected=1296.0), p = 0.91.

Table 4.2

Wilcoxon Rank Sums for the Renaissance STAR Score

Group	N	Sum of Scores	Expected Under H_0	Std Dev Under H_0	Mean Score
Control	36	1306.0	1296.0	86.94	36.28
Treatment	35	1250.0	1260.0	86.94	35.71

Note: $p = 0.91 > 0.05$ (Appendix D, p.173)

The results of the Wilcoxon Rank Sums test reaffirm the results of the parametric t-test that the null hypothesis fails to be rejected and conclude that there is no statistically significant difference between the scores of the Control and Treatment groups on the Renaissance STAR pre-test despite the Treatment groups' mean score being more than one and one-half points higher than the Control groups' mean score.

Teacher-generated pre-test and post-test

Scores from the teacher-generated pre-test and post-test were collected for assessments given prior to beginning and at the conclusion of the functions unit. In the first phase, the pre-test scores were used to determine a baseline achievement level in pre-algebra, specifically on functions to establish whether the Treatment and Control groups had similar achievement in functions prior to the intervention. Scores on the pre-test for the Treatment and Control groups were compared using an independent-samples t-test. This test compared the pre-algebra functions scores of students enrolled in the DAIP activity and students enrolled in traditional pre-algebra. The results are reflected in Table 4.3 and indicate that students in the Treatment group had a marginally higher score ($M = 0.26$, $SD = 0.12$) than students in the Control group

(M = 0.23, SD = 0.13), p = 0.17.

Table 4.3

Teacher-generated Pre-test Results

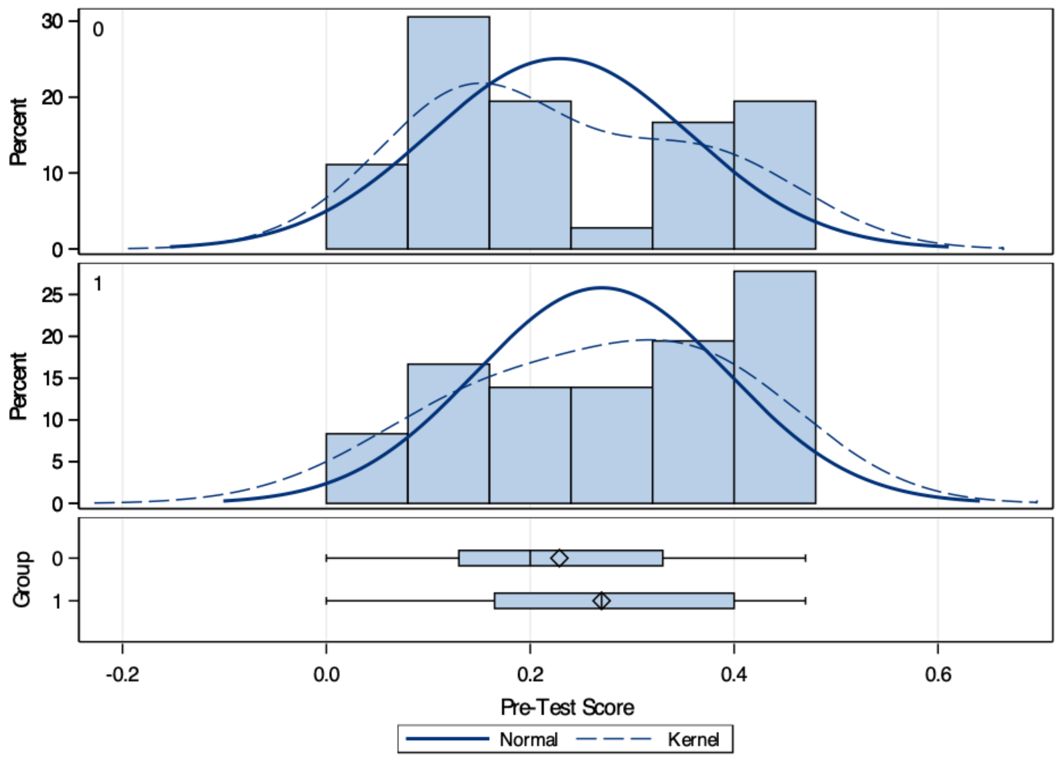
Group	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	36	0.23	0.13	0.02	0	0.47
Treatment	36	0.26	0.12	0.02	0	0.47

Note: $p = 0.17 > 0.05$ (Appendix E, p. 190)

The data was not normally distributed as indicated in Figure 4.2. The box plots below the histogram indicate that the two groups significantly overlap with the mean of the Treatment group slightly higher than that of the Control group.

Figure 4.2

Data Distribution for Teacher-generated Pretest



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. (Appendix E, p. 191)

Because the data did not appear to be normally distributed, a parametric test does not have adequate power to accurately predict significance and a non-parametric Wilcoxon Rank Sums test was also administered. (Table 4.4)

Table 4.4

Wilcoxon Rank Sums for the Pre-Test Score

Group	N	Sum of Scores	Expected Under H_0	Std Dev Under H_0	Mean Score
Control	36	1190.50	1314.0	87.43	33.07
Treatment	36	1437.50	1314.0	87.43	39.93

Note: $p = 0.16 > 0.05$ (Appendix E, p. 192)

The results for the Wilcoxon Rank Sums test reflected in Table 4.4 indicate that the Sum of Scores for the Treatment group (SoS=1437.50, Expected = 1314.0) is slightly higher than the Control group (SoS=1190.50, Expected=1314.0), p = 0.16. Despite the Treatment groups' mean score being more than three and one-half percentage points higher than the Control groups' mean score (Table 4.3), the results of the Wilcoxon test reaffirm the results of the parametric t-test (Table 4.4) and indicate that the null hypothesis fails to be rejected and there is no statistically significant difference between the scores of the Control and Treatment groups on the teacher-generated pre-test.

Attitudes Toward Mathematics Inventory

The Attitudes Toward Mathematics Inventory (ATMI) developed by Martha Tapia from the University of Alabama (Tapia, 1996) consists of 40 questions designed to measure attitudes toward mathematics in high school and college students. The questions were in not altered to accommodate the middle school student population. The forty questions address four key subscales that pertain to attitudes and confidence in learning mathematics. Survey questions 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 & 40 address student self-confidence. Questions 1, 2, 4, 5, 6, 7, 8, 35, 36 & 39 measure perceived value in studying mathematics. Questions 3, 24, 25, 26, 27, 29, 30, 31, 37 & 38 determine student enjoyment of mathematics. Finally, questions 23, 28, 32, 33 & 34 evaluate student motivation to learn mathematics. Responses to the questions were compiled into the four subscales and an average, composite score was generated for each respondent for each subscale.

In the first phase, the ATMI was administered to all students in both the treatment and control groups prior to the teacher-generated pre-test. The same survey was administered before the teacher-generated post-test at the end of the functions unit. These responses from the ATMI were collected for surveys given prior to pre-test and post-test after the functions unit.

In the first phase, composite scores on the pre-unit survey for the treatment and control groups were compared using an independent-samples t-test with the intent to establish whether the treatment and control groups had similar attitudes toward mathematics prior to the intervention. This test compared the attitudes towards mathematics of students enrolled in the DAIP activity with students enrolled in traditional pre-algebra. (Table 4.5)

Table 4.5

ATMI Pre-survey Result Summary

	Self-confidence		Value		Enjoyment		Motivation	
Group	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Control	3.11	0.40	4.02	0.56	3.64	0.58	3.65	0.66
Treatment	2.97	0.37	3.92	0.73	3.50	0.71	3.39	0.96
p-value	p = 0.12 > 0.05		p = 0.52 > 0.05		p = 0.34 > 0.05		p = 0.18 > 0.05	

Note: p-values indicate no statistical difference between groups for any subscale.

The following tables provide expanded data for the summary of the Self-confidence Subscale shown in Table 4.5. (Appendices F, G, H, J). The results of the pre-survey two-sample t-test are reflected in Table 4.6 and indicate that students in the Treatment group had a marginally lower score (M = 2.97, SD = 0.37) than students in the Control group (M = 3.11, SD = 0.40), p = 0.13.

Table 4.6

Pre-survey Results for Self-confidence Subscale

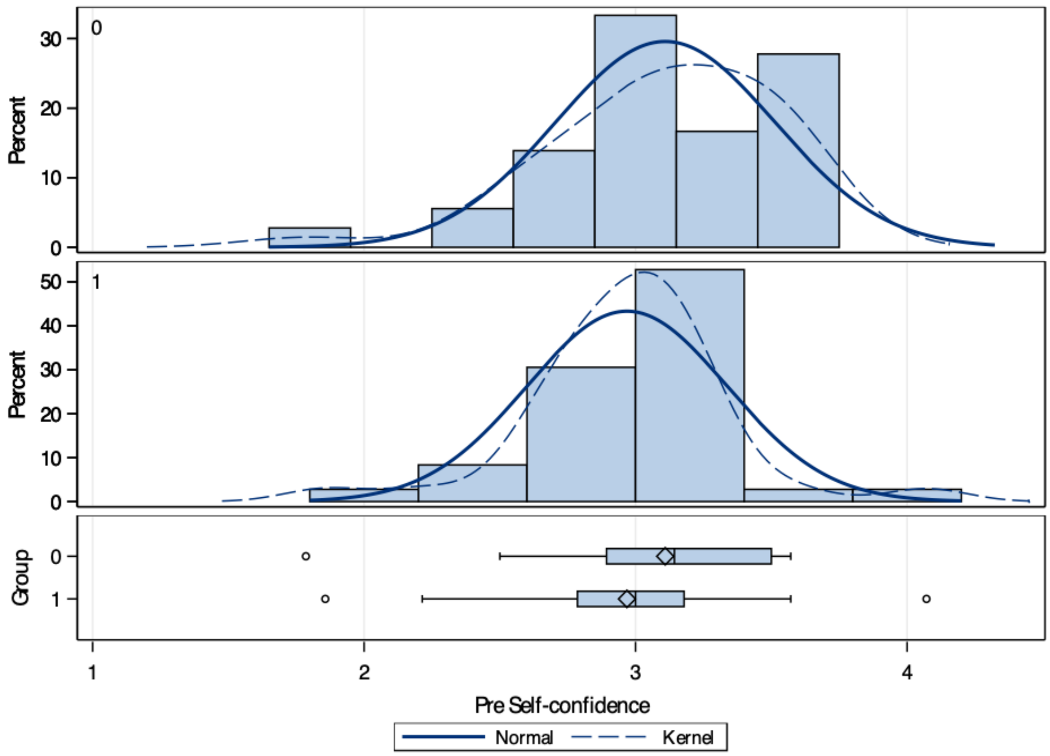
Group	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	36	3.11	0.40	0.07	1.79	3.57
Treatment	36	2.97	0.37	0.06	1.86	4.07

Note: Pooled p = 0.13 > 0.05 Satterthwaite p = 0.13 > 0.05 (Appendix F, p. 209)

The data was not normally distributed as indicated in Figure 4.3. The box plots below the histogram indicate that the two groups significantly overlap with the mean of the Control group somewhat higher than that of the Treatment group.

Figure 4.3

Data Distribution for Pre-survey Self-confidence Subscale



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. (Appendix F, p. 210)

Because the data did not appear to be normally distributed, a parametric t-test does not have adequate power to accurately predict significance and a non-parametric Wilcoxon Rank Sums test was also administered. (Table 4.7)

Table 4.7

Wilcoxon Rank Sums for Pre-survey Results for Self-confidence Subscale

Group	N	Sum of Scores	Expected Under H_0	Std Dev Under H_0	Mean Score
Control	36	1482.50	1314.0	88.47	41.18
Treatment	36	1145.50	1314.0	88.47	31.82

Note: p = 0.06 > 0.05 (Appendix F, p. 211)

The results for the Wilcoxon Rank Sums test reflected in Table 4.7 indicate that the Sum of Scores for the Treatment group (SoS=1145.50, Expected = 1314.0) is slightly higher than the Control group (SoS=1482.50, Expected=1314.0), p = 0.06. The results of the Wilcoxon test reaffirm the results of the parametric t-test and indicate that the null hypothesis fails to be rejected and conclude that there is no statistically significant difference between the responses of the Control and Treatment groups on the Self-confidence Subscale of the pre-survey.

The following tables provide expanded data for the summary of the Value Subscale shown in Table 4.5.

Table 4.8

Pre-survey Results for Value Subscale

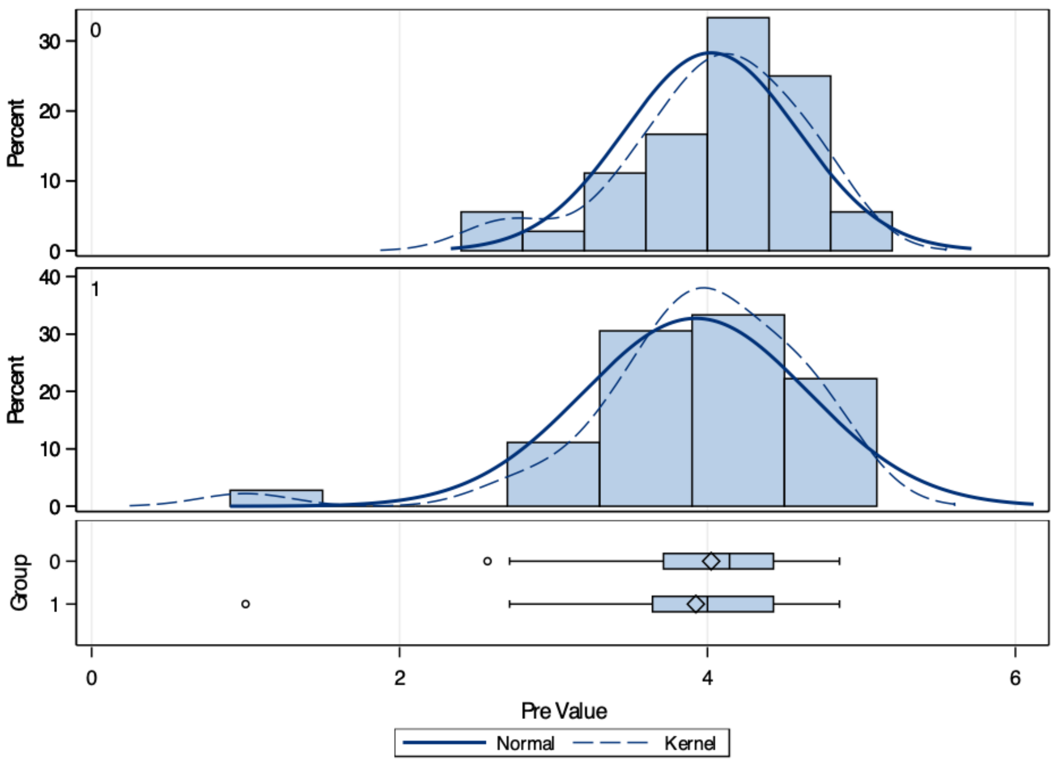
Group	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	36	4.02	0.56	0.09	2.57	4.86
Treatment	36	3.92	0.73	0.12	1.00	4.86

Note: Pooled p = 0.52 > 0.05 Satterthwaite p = 0.52 > 0.05 (Appendix G, p. 227)

The results of the pre-survey two-sample t-test are reflected in Table 4.8 and indicate that students in the Treatment group had a marginally lower score (M = 3.92, SD = 0.73) than students in the Control group (M = 4.02, SD = 0.56), p = 0.52. The data was not normally distributed as indicated in Figure 4.4. The box plots below the histogram indicate that the two groups significantly overlap with the mean of the Control group slightly higher than that of the Treatment group.

Figure 4.4

Data Distribution for Pre-survey Value Subscale



Note: Group 0 represents the Control group, Group 1 represents the Treatment group.

(Appendix G, p. 228)

Because the data did not appear to be normally distributed, a parametric test does not have adequate power to accurately predict significance, and a non-parametric Wilcoxon Rank Sums test was also administered. (Table 4.9)

Table 4.9

Wilcoxon Rank Sums for Pre-survey Results for Value Subscale

Group	N	Sum of Scores	Expected Under H_0	Std Dev Under H_0	Mean Score
Control	36	1362.0	1314.0	88.45	37.83
Treatment	36	1266.0	1314.0	88.45	35.17

Note: $p = 0.59 > 0.05$ (Appendix G, p. 229)

The results for the Wilcoxon Rank Sums test reflected in Table 4.9 indicate that the Sum of Scores for the Treatment group (SoS=1266.0, Expected = 1314.0) is slightly higher than the Control group (SoS=1362.0, Expected=1314.0), $p = 0.59$. The results of the Wilcoxon test reaffirm the results of the parametric t-test and indicate that the null hypothesis fails to be rejected and conclude that there is no statistically significant difference between the responses of the Control and Treatment groups on the Value Subscale of the pre-survey.

The following tables provide expanded data for the summary of the Enjoyment Subscale shown in Table 4.5.

Table 4.10

Pre-survey Results for Enjoyment Subscale

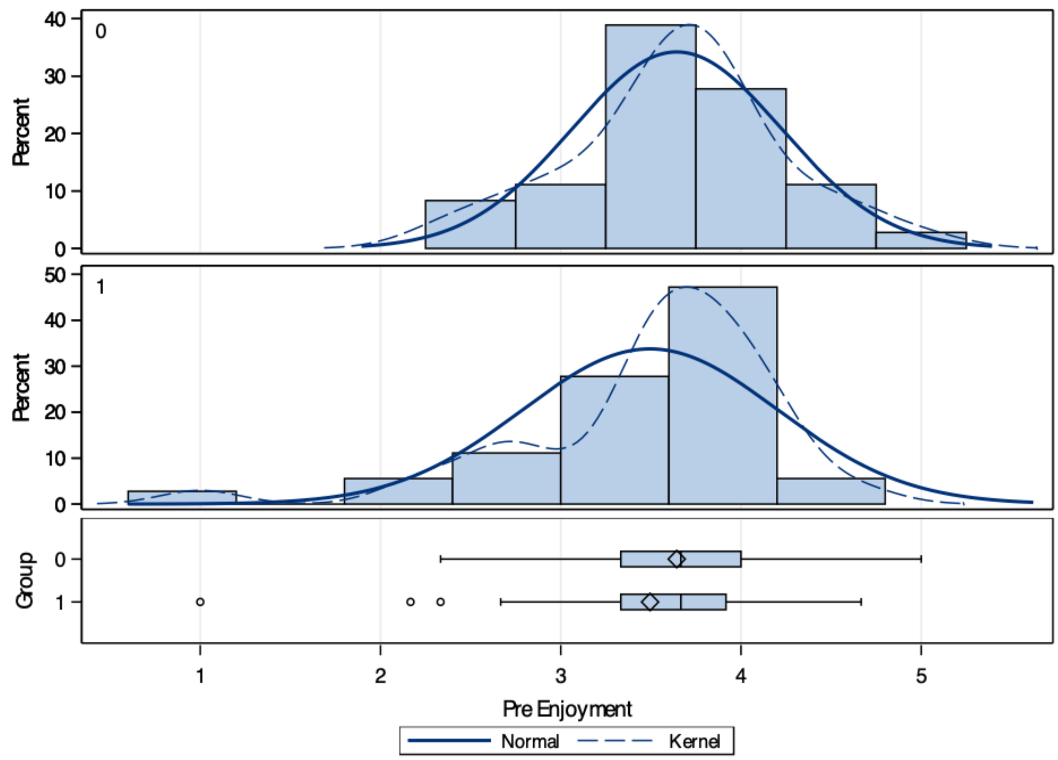
Group	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	36	3.64	0.58	0.10	2.33	5.00
Treatment	36	3.50	0.71	0.12	1.00	4.67

Note: Pooled p = 0.34 > 0.05 Satterthwaite p = 0.34 > 0.05 (Appendix H, p. 246)

The results of the pre-survey two-sample t-test are reflected in Table 4.10 and indicate that students in the Treatment group had a marginally lower score (M = 3.50, SD = 0.71) than students in the Control group (M = 3.64, SD = 0.58), p = 0.34. The data was not normally distributed as indicated in Figure 4.5. The box plots below the histogram indicate that the two groups significantly overlap with the mean of the Control group slightly higher than that of the Treatment group.

Figure 4.5

Data Distribution for Pre-survey Enjoyment Subscale



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. (Appendix H, p. 247)

Because the data did not appear to be normally distributed, a parametric test does not have adequate power to accurately predict significance and a non-parametric Wilcoxon Rank Sums test was also administered. (Table 4.11)

Table 4.11

Wilcoxon Rank Sums for Pre-survey Results for Enjoyment Subscale

Group	N	Sum of Scores	Expected Under H_0	Std Dev Under H_0	Mean Score
Control	36	1357.50	1314.0	88.22	37.71
Treatment	36	1270.50	1314.0	88.22	35.29

Note: p = 0.59 > 0.05 (Appendix H, p. 248)

The results for the Wilcoxon Rank Sums test reflected in Table 4.11 indicate that the Sum of Scores for the Treatment group (SoS=1270.50, Expected = 1314.0) is slightly higher than the Control group (SoS=1357.50, Expected=1314.0), p = 0.59. The results of the Wilcoxon test reaffirm the results of the parametric t-test and indicate that the null hypothesis fails to be rejected and conclude that there is no statistically significant difference between the responses of the Control and Treatment groups on the Enjoyment Subscale of the pre-survey.

The following tables provide expanded data for the summary of the Motivation Subscale shown in Table 4.5.

Table 4.12

Pre-survey Results for Motivation Subscale

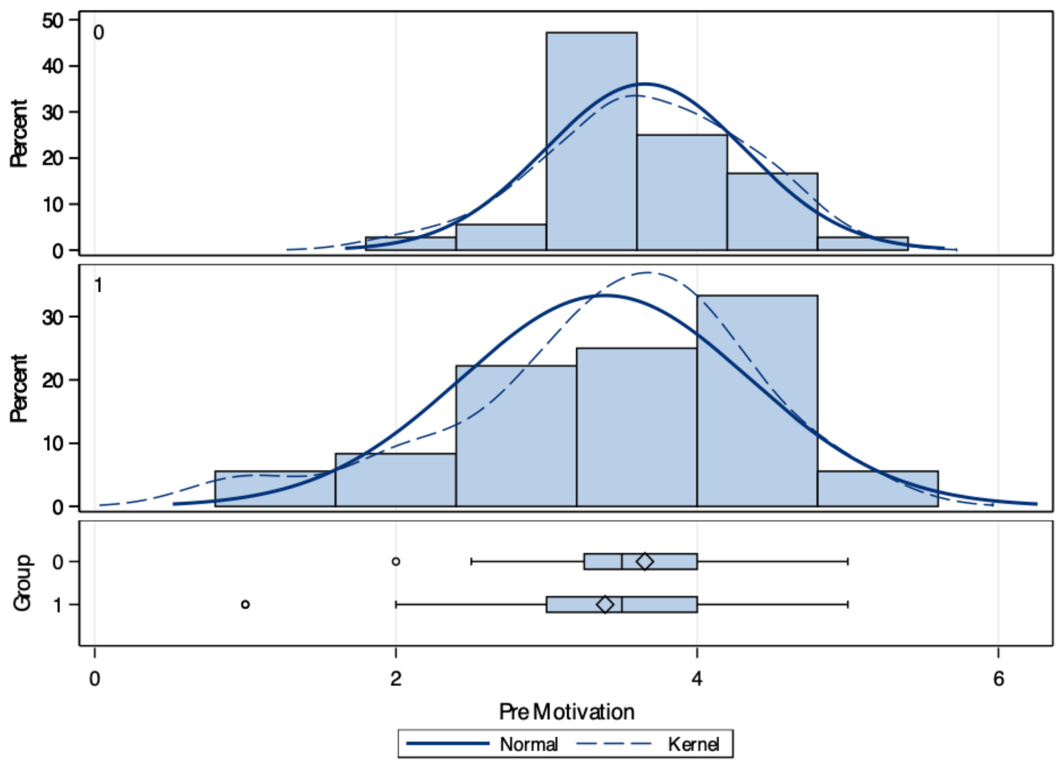
Group	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	36	3.65	0.66	0.11	2.00	5.00
Treatment	36	3.39	0.956	0.16	1.00	5.00

Note: Pooled p = 0.18 > 0.05 Satterthwaite p = 0.18 > 0.05 (Appendix J, p. 265)

The results of the pre-survey two-sample t-test are reflected in Table 4.12 and indicate that students in the Treatment group had a marginally lower score (M = 3.39, SD = 0.96) than students in the Control group (M = 3.65, SD = 0.66), p =0.18. The data was not normally distributed as indicated in Figure 4.6.

Figure 4.6

Data Distribution for Pre-survey Motivation Subscale



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. (Appendix H, p. 266)

Because the data did not appear to be normally distributed, a parametric test does not have adequate power to accurately predict significance and a non-parametric Wilcoxon Rank Sums test was also administered. (Table 4.13)

Table 4.13

Wilcoxon Rank Sums for Pre-survey Results for Motivation Subscale

Group	N	Sum of Scores	Expected Under H_0	Std Dev Under H_0	Mean Score
Control	36	1402.50	1314.0	86.82	38.96
Treatment	36	1225.50	1314.0	86.82	34.04

Note: $p = 0.31 > 0.05$ (Appendix H, p. 267)

The results for the Wilcoxon Rank Sums test reflected in Table 4.13 indicate that the Sum of Scores for the Treatment group (SoS=1225.50, Expected = 1314.0) is slightly higher than the Control group (SoS=1402.50, Expected=1314.0), $p = 0.31$. The results of the Wilcoxon test reaffirm the results of the parametric t-test and indicate that the null hypothesis fails to be rejected and conclude that there is no statistically significant difference between the responses of the Control and Treatment groups on the Motivation Subscale of the pre-survey.

Results

To test the first hypothesis addressing research question one, the results from the Renaissance STAR Exam standardized test and a teacher-generated post-test were analyzed and compared to pre-test scores for both the Treatment and Control groups.

H_01: There is no significant difference in the achievement in pre-algebra as measured by the Renaissance STAR Benchmark Assessment (Renaissance, 1998) and a teacher-generated pre- and post-assessment between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

H_a1: There is a significant difference in the achievement in pre-algebra as measured by the Renaissance STAR Benchmark Assessment (Renaissance, 1998) and a teacher-generated pre- and post-assessment between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

Achievement in Pre-Algebra: Renaissance STAR Exam

In the second phase, composite scores from the Treatment and Control groups' September and January Renaissance STAR Exam were then compared. To determine if a statistically significant difference in achievement existed between the groups, scores from the pre-test and post-test Renaissance STAR Exam were compared using an Analysis of Covariance (ANCOVA) between the Treatment and Control groups. (Table 4.14)

Table 4.14

ANCOVA results for Renaissance STAR Exam

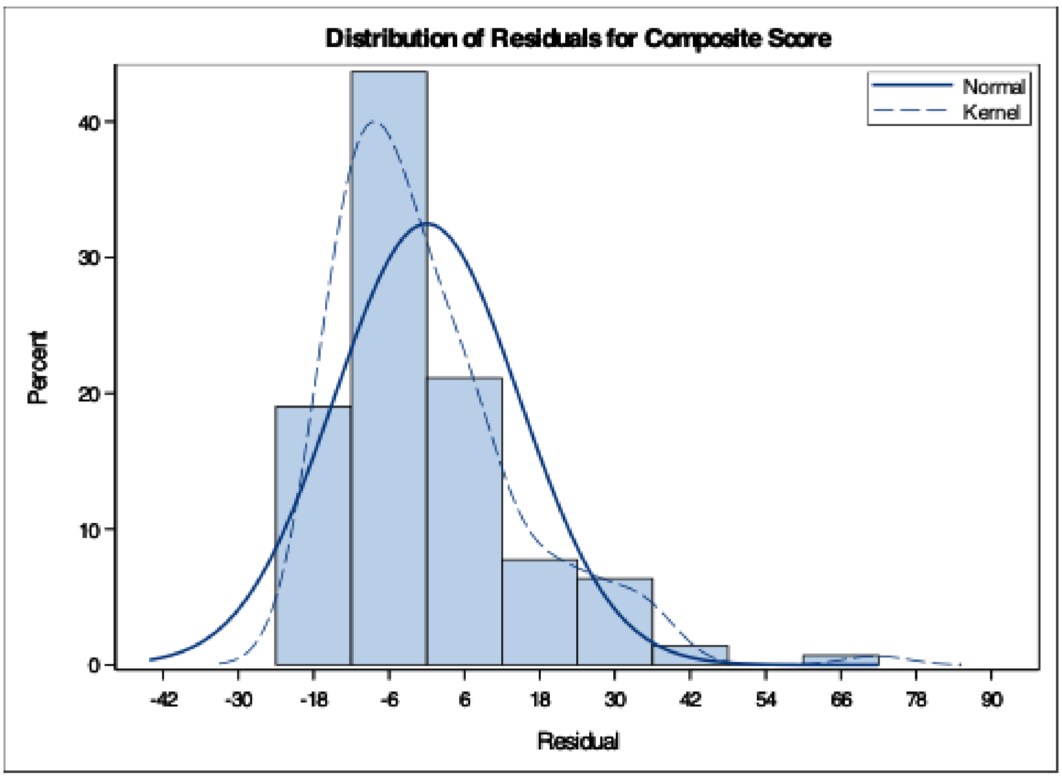
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	972.14	324.05	1.46	0.23
Error	138	30634.58	221.99		
Corrected Total	141	31606.72			

Note: The results of the ANCOVA indicate no statistically significant difference in assessment scores between the Treatment and Control groups. (Appendix D, p. 174)

Table 4.14 reflects the results of the ANCOVA which show that experiencing the DAIP activity had little significant, positive effect on the Renaissance STAR exam post-test scores when controlling for Renaissance STAR exam pre-test scores, F = 1.46, $p < 0.23$. The histogram (Fig 4.7) from the ANCOVA analyzing the Renaissance STAR Exam indicates that the residuals were not normally distributed.

Figure 4.7

Distribution of Residuals on the Renaissance STAR Exam



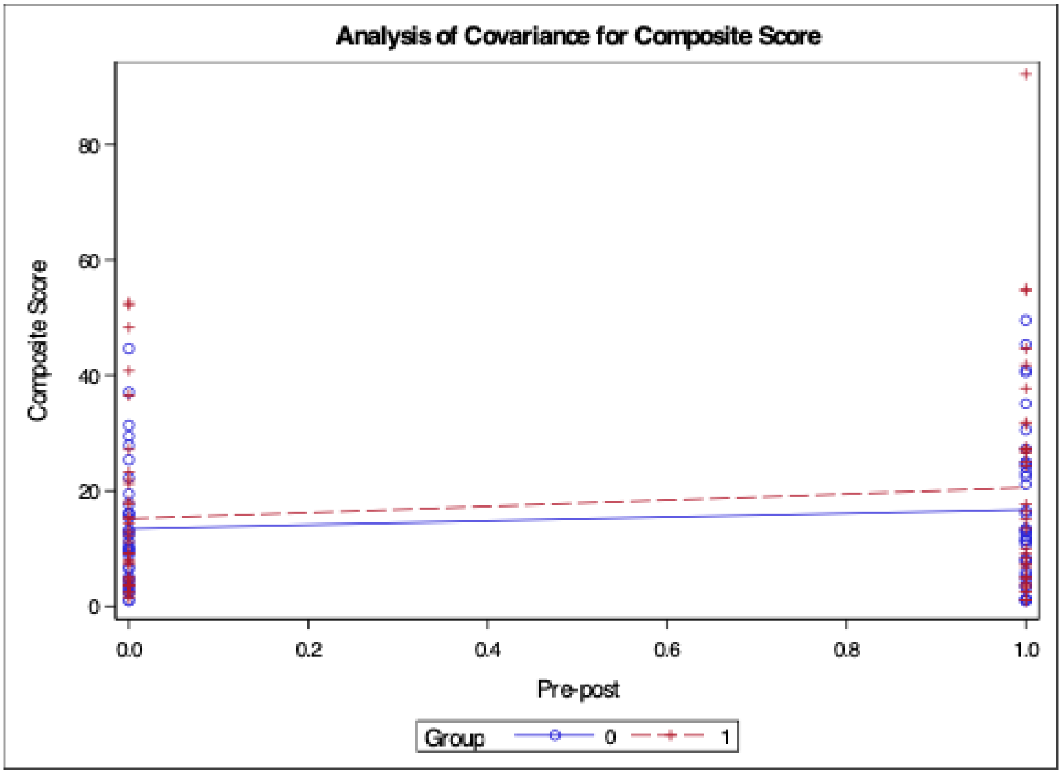
Note: (Appendix D, p. 178)

Despite the residuals not being normally distributed, the ANCOVA indicates similar improvement on post-test assessment scores in each group over their pre-test scores. The interaction plot (Fig 4.8) comparing the change in the two groups reveals that while the

Treatment group scored higher, the relative difference in improvement was not statistically significant.

Figure 4.8

Interaction Plot Comparing Groups on the Renaissance STAR Exam



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. Pre-Post values indicate scores on the pre-test (0) and post-test (1). (Appendix D, p. 179)

Since the residuals were not normally distributed, the ANCOVA does not have adequate power to confirm a statistical significance between the Treatment and Control groups. A two-sample t-test comparing post-test scores was administered to confirm the results of the ANCOVA. (Table 4.15) The results are reflected in Table 4.15 and indicate that students in the Treatment group had a marginally higher score (M = 20.53, SD = 19.81) than students in the Control group (M = 16.75, SD = 13.21), p = 0.35.

Table 4.15

Renaissance STAR Exam Post-test t-test Results

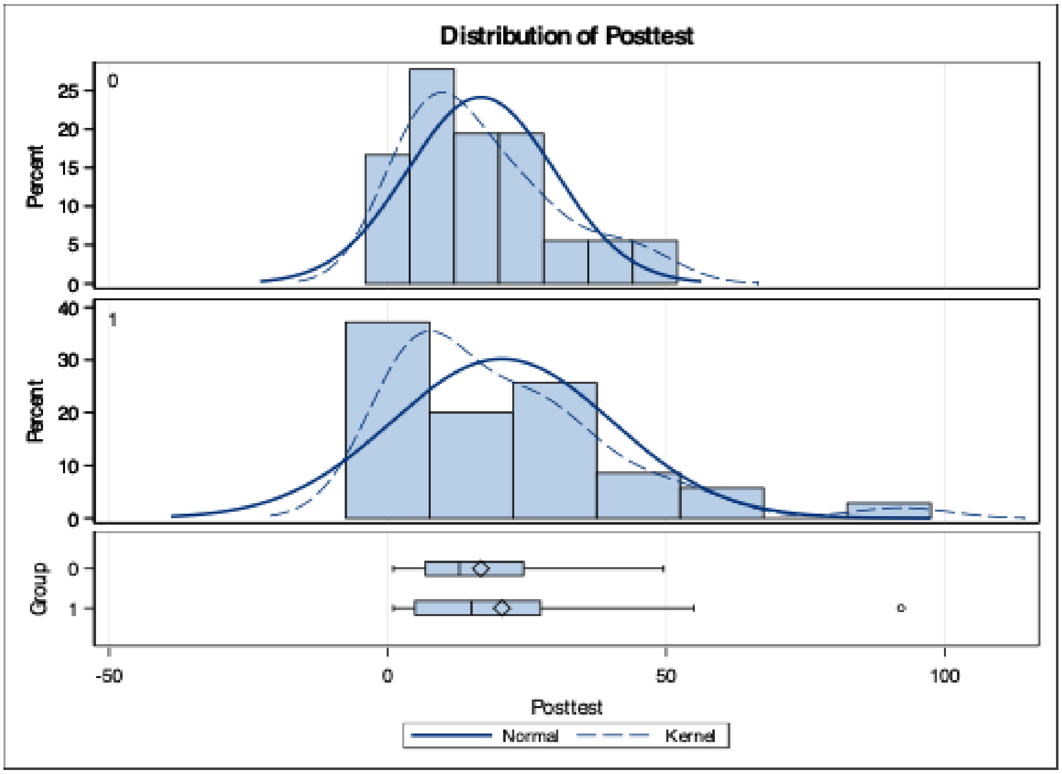
Group	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	36	16.75	13.21	2.20	1.00	49.55
Treatment	35	20.53	19.81	3.35	1.00	92.18

Note: Pooled $p = 0.35 > 0.05$ Satterthwaite $p = 0.35 > 0.05$ (Appendix D, p. 187)

The unusual Maximum value of 92.18 for the treatment group was due to a single subject's anomalous score. The same subject returned a more expected score on the teacher-generated assessment, therefore the data was not treated as an outlier. The data was not normally distributed as indicated in Figure 4.9. The box plots below the histogram indicate that the two groups significantly overlap with the mean of the Treatment group slightly higher than that of the Control group.

Figure 4.9

Data Distribution for the Renaissance STAR Exam Post-test t-test



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. (Appendix D, p. 188)

Because the data did not appear to be normally distributed, a parametric test does not have adequate power to accurately predict significance and a non-parametric Wilcoxon Rank Sums test was also administered. (Table 4.16)

Table 4.16

Wilcoxon Rank Sums Results for the Renaissance STAR Exam Post-test t-test

Group	N	Sum of Scores	Expected Under H_0	Std Dev Under H_0	Mean Score
Control	36	1256.0	1296.0	86.93	34.89
Treatment	35	1300.0	1260.0	86.93	37.143

Note: p = 0.65 > 0.05 (Appendix D, p. 189)

The results for the Wilcoxon Rank Sums test reflected in Table 4.16 indicate that the Sum of Scores for the Treatment group (SoS=1300.0, Expected = 1260.0) is slightly higher than the Control group (SoS=1256.0, Expected=1296.0), p = 0.65. Since the residuals on the ANCOVA were not normally distributed, pre-test and post-test scores were compared using paired t-tests for both Treatment and Control groups to confirm the relative change between groups. (Table 4.17)

Table 4.17

Paired T-test Results for the Renaissance STAR Exam Pre-test and Post-test

	N	Mean	Std Dev	Std Err	Minimum	Maximum	p-value
Control	36	3.31	8.26	1.38	-14.82	21.64	0.02
Treatment	35	5.38	17.24	2.91	-24.91	87.36	0.07

Note: The negative minimum values indicate that some students scored lower on the post-test than the pre-test. (Appendix D, p. 190)

The results are reflected in Table 4.17 and indicate that students in the Treatment group had a marginally higher score (M = 5.38, SD = 17.24) than students in the Control group

(M = 3.31, SD = 8.26). The results of the Wilcoxon test and the paired t-tests reaffirm the results of the ANCOVA and indicate that the null hypothesis fails to be rejected and conclude that there is no statistically significant difference in improvement between the scores of the Control and Treatment groups on the Renaissance STAR Exam post-test, despite the Treatment groups' mean score being more than two points higher than the Control groups' mean score.

Achievement in Pre-Algebra: Teacher-generated pre-test and post-test

All 72 students who agreed to participate in the research were included in the following data. In the second phase, both the Treatment and Control groups received identical teacher-generated post-test following the unit on functions. Scores from the pre-test and post-test teacher-generated assessment were then compared using an Analysis of Covariance (ANCOVA) between the Treatment and Control groups' pre-test and post-test composite scores to determine if a statistically significant difference in achievement existed between the groups. (Table 4.18)

Table 4.18

ANCOVA results for Teacher-generated Assessment

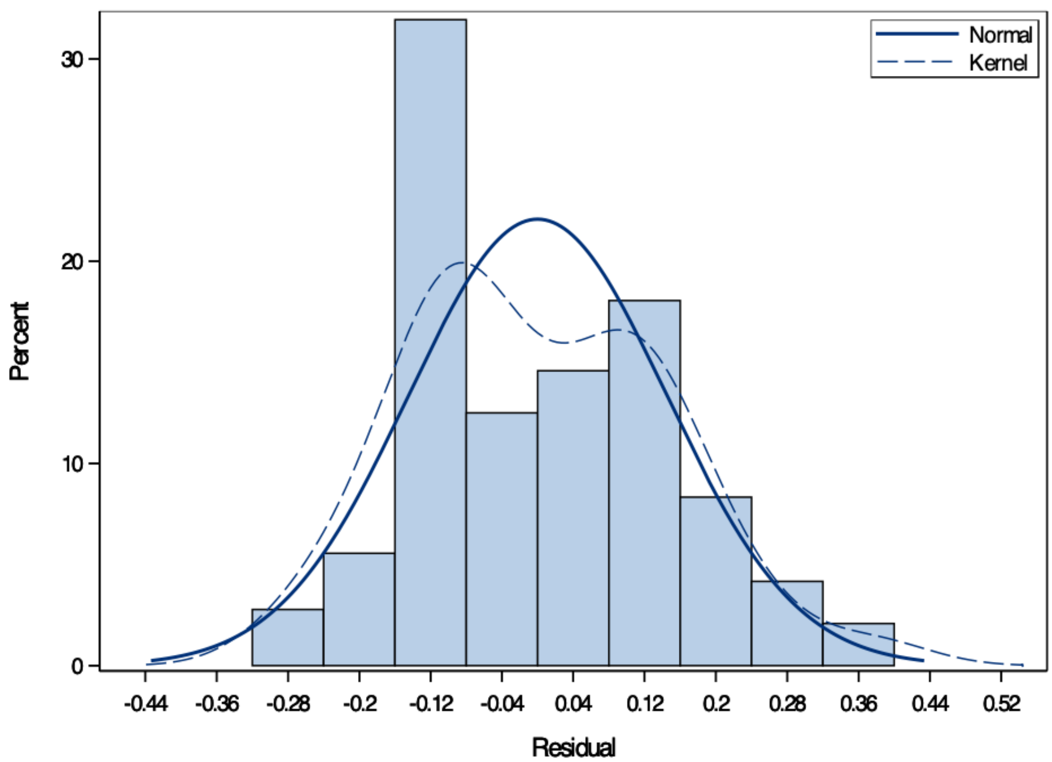
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.75	0.25	11.68	<.0001
Error	140	2.99	0.021		
Corrected Total	143	3.74			

Note: The results of the ANCOVA indicate a statistically significant difference in assessment scores between the Treatment and Control groups. (Appendix E, p. 193)

Table 4.18 reflects the results of the ANCOVA which show that experiencing the DAIP activity had a significant, positive effect on the Teacher-generated assessment post-test scores when controlling for Teacher-generated assessment pre-test scores, $F = 11.68$, $p < 0.0001$. Despite indicating a statistically significant difference between the Treatment and Control groups' scores on the Teacher-generated Assessment, the residuals were not normally distributed. (Fig 4.10)

Figure 4.10

Distribution of Residuals on the Teacher-generated Assessment

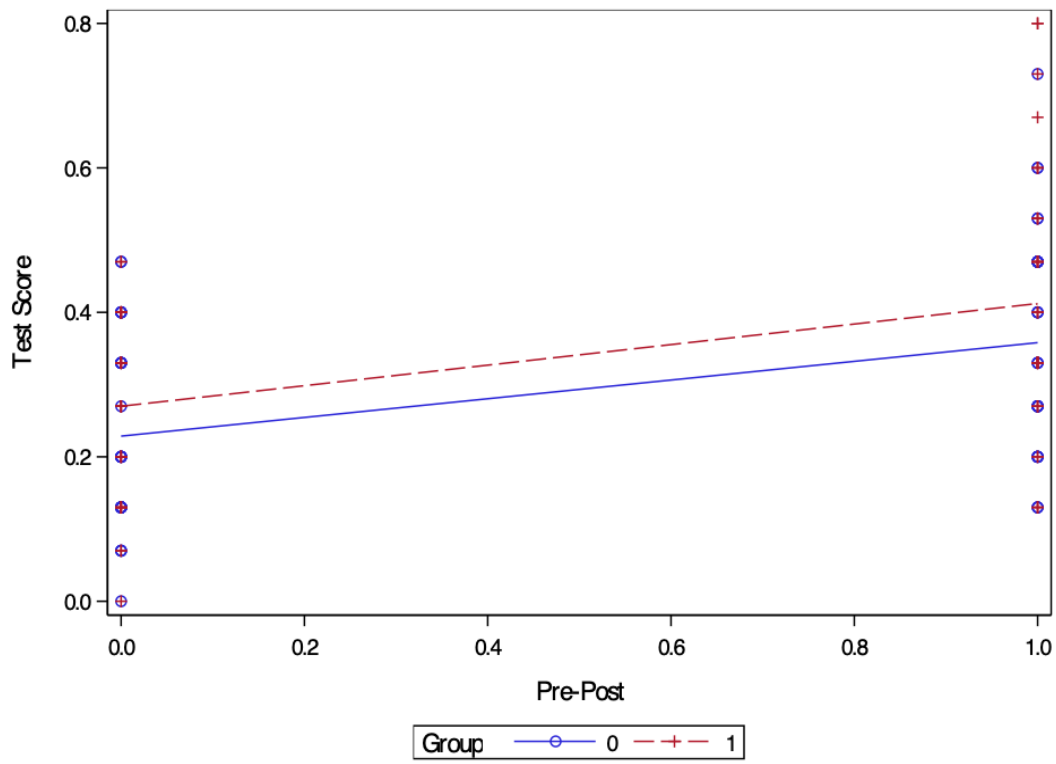


Note: (Appendix E, p. 197)

While the residuals were not normally distributed, the ANCOVA indicates similar improvement on post-test scores in each group over their pre-test scores on the teacher-generated assessment. The interaction plot (Fig 4.11) comparing the change in the two groups reveals that although both groups improved by a statistically significant amount, the Treatment group scored higher, however, the relative difference in improvement between groups appears to not be statistically significant.

Figure 4.11

Interaction Plot Comparing Groups on the Teacher-generated Assessment



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. Pre-Post values indicate scores on the pre-test (0) and post-test (1). (Appendix E, p. 198)

Since the residuals were not normally distributed, the ANCOVA does not have adequate power to confirm a statistical significance between the Treatment and Control groups. Paired t-tests comparing pre-test and post-test scores were administered to confirm the results of the ANCOVA. (Table 4.19)

Table 4.19

Paired T-test Results for the Teacher-generated Pre-test and Post-test

	N	Mean	Std Dev	Std Err	Minimum	Maximum	p-value
Control	36	0.13	0.16	0.03	- 0.20	0.40	<.0001
Treatment	36	0.14	0.20	0.03	-0.20	0.67	0.0002

Note: The negative minimum values indicate that some students scored lower on the post-test than the pre-test. (Appendix E, p. 200)

The results are reflected in Table 4.18 and indicate that students in the Treatment group had a marginally higher score (M = 0.14, SD = 0.20) than students in the Control group (M = 0.13, SD = 0.16). Improvement by both groups appears to be statistically significant as indicated by the p-values, however, because the residuals were not normally distributed and statistical significance between the Treatment and Control groups could not be established with the ANCOVA, a two-sample t-test comparing post-test scores was administered to confirm the results. (Table 4.20)

Table 4.20

Teacher-generated Post-test Results

Group	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	36	0.36	0.15	0.03	0.13	0.73
Treatment	36	0.41	0.18	0.03	0.13	0.80

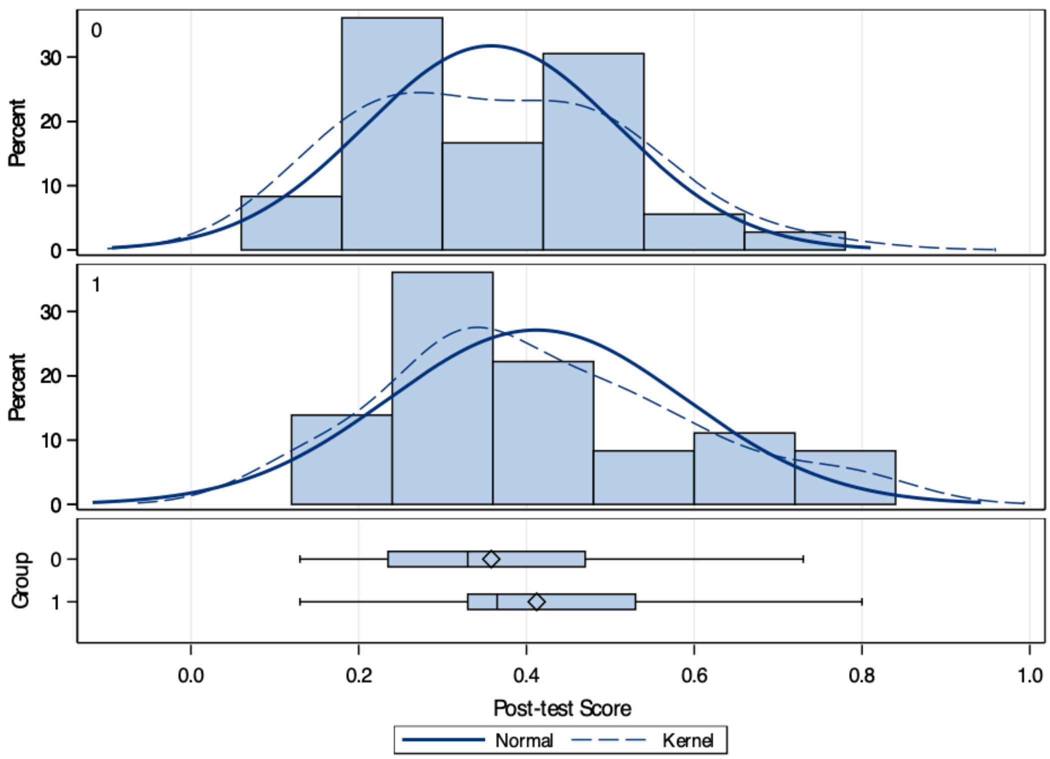
Note: Pooled p = 0.17> 0.05 Satterthwaite p = 0.17 > 0.05 (Appendix E, p. 206)

The results are reflected in Table 4.19 and indicate that students in the Treatment group had a marginally higher score (M = 0.41, SD = 0.18) than students in the Control group

(M = 0.36, SD = 0.15), p = 0.16. The data was not normally distributed as indicated in Figure 4.12. The box plots below the histogram indicate that the two groups significantly overlap with the mean of the Treatment group slightly higher than that of the Control group.

Figure 4.12

Data Distribution for Teacher-generated Post-test



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. (Appendix E, p. 207)

Because the data did not appear to be normally distributed, a parametric t-test does not have adequate power to accurately predict significance and a non-parametric Wilcoxon Rank Sums test was also administered. (Table 4.21)

Table 4.21

Wilcoxon Rank Sums for Post-test Results for Teacher-generated Assessment

Group	N	Sum of Scores	Expected Under H_0	Std Dev Under H_0	Mean Score
Control	36	1198.50	1314.0	87.95	33.29
Treatment	36	1429.50	1314.0	87.95	39.71

Note: p = 0.20 > 0.05 (Appendix E, p. 208)

The results for the Wilcoxon Rank Sums test reflected in Table 4.21 indicate that the Sum of Scores for the Treatment group (SoS=1429.50, Expected = 1314.0) is slightly higher than the Control group (SoS=1198.50, Expected=1314.0), p = 0.20. The results of the Wilcoxon test and the paired t-tests reaffirm the results of the parametric t-test and indicate that the null hypothesis fails to be rejected and conclude that there is no statistically significant difference in improvement between the scores of the Control and Treatment groups on the teacher-generated post-test, despite the Treatment groups' mean score being more than five percentage points higher than the Control groups' mean score.

Results on the Renaissance STAR exam and the teacher-generated assessment after the implementation of the DAIP activity indicate that there is no statistically significant difference between the Control and Treatment groups' performance. While both groups' scores improved, the relative improvement between the groups indicated no statistically significant difference. (Table 4.22)

Table 4.22

Change in Score Summary on the Renaissance STAR and Teacher-generated Assessments

Pre-Assessment Scores

	Renaissance STAR Exam			Teacher-generated Assessment		
Group	N	Mean	Std Dev	N	Mean	Std Dev
Control	36	13.44	10.54	36	0.23	0.13
Treatment	35	15.15	14.65	36	0.26	0.12
p-value	p = 0.57 > 0.05			p = 0.16 > 0.05		

Post-Assessment Scores

	Renaissance STAR Exam			Teacher-generated Assessment		
Group	N	Mean	Std Dev	N	Mean	Std Dev
Control	36	16.75	13.21	36	0.36	0.15
Treatment	35	20.53	19.81	36	0.41	0.18
p-value	p = 0.35 > 0.05			p = 0.17 > 0.05		

Attitudes Toward Mathematics Inventory

All 72 students who agreed to participate in the research were included in the following data. In the second phase, composite scores from the pre-unit survey and the post-unit survey were then compared using an Analysis of Covariance (ANCOVA) between the Treatment and Control groups' pre-unit survey and post-unit survey composite scores to determine if a statistically significant difference in attitudes and confidence in learning mathematics existed between the groups.

To test hypothesis two addressing the second research question, the results from the ATMI survey responses were analyzed and compared to pre-unit survey scores for both the Treatment and Control groups.

$H_0$2: There is no significant difference between the confidence and interest in STEM topics as measured by scores on the Attitudes Toward Mathematics Inventory (ATMI) (Tapia, 1996) between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

H_a2: There is a significant difference between the confidence and interest in STEM topics as measured by scores on the Attitudes Toward Mathematics Inventory (ATMI) (Tapia, 1996) between disenfranchised African American middle school students experiencing a project-based learning activity using design-thinking and expressed through an artform (DAIP) and those students experiencing a traditional pre-algebra curriculum.

The subscale evaluating students' self-confidence in their ability to solve mathematics problems was evaluated using an ANCOVA which compared post-survey results to the students' responses to the same survey administered prior to the intervention. This test indicated the relative change in scores between the two groups. (Table 4.23)

Table 4.23

ANCOVA results for the Post-survey Self-confidence Subscale

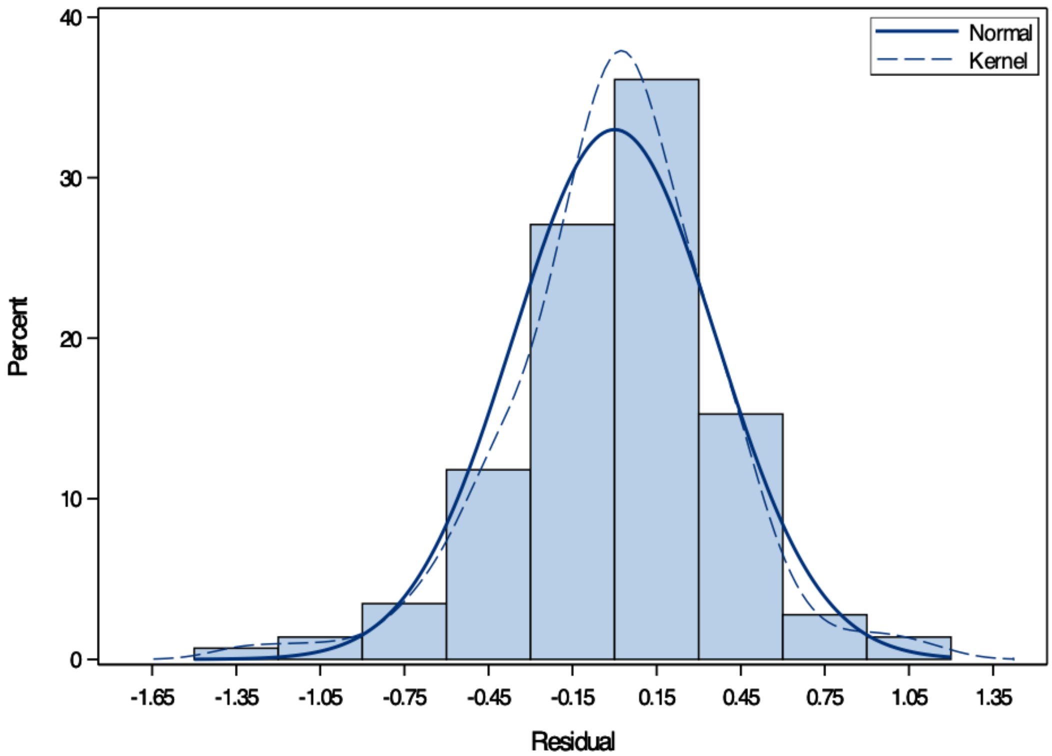
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.37	0.123	0.91	0.44
Error	140	18.81	0.13		
Corrected Total	143	19.19			

Note: The results of the ANCOVA indicate no statistically significant difference in survey responses between the Treatment and Control groups. (Appendix F, p. 212)

Table 4.23 reflects the results of the ANCOVA which show that experiencing the DAIP activity had no significant, positive effect on the students' Self-confidence subscale post-survey scores when controlling for survey Self-confidence subscale pre-survey scores, $F = 0.91$, $p < 0.44$. The histogram of the distribution of residuals on the ANCOVA analyzing the post-survey Self-confidence subscale indicates a reasonably normal distribution, therefore the ANCOVA has adequate power allowing for the results to be accepted. (Fig 4.13)

Figure 4.13

Distribution of Residuals the Post-survey Self-confidence Subscale

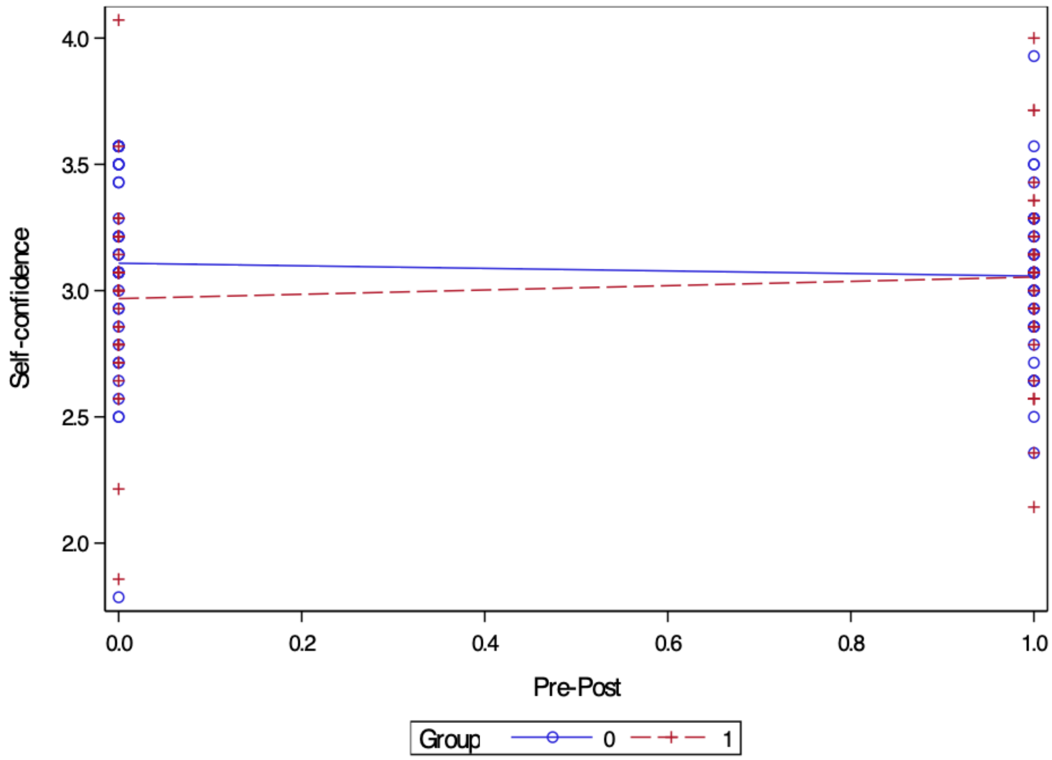


Note: (Appendix F, p. 216)

The Interaction Plot for the Self-confidence Subscale (Fig 4.14) compares the change in the two groups and indicates that following the intervention, both Control and Treatment groups experienced similar levels of self-confidence in their ability to solve mathematics problems, although the Treatment groups' self-confidence appears to have increased while the Control groups' self-confidence slightly decreased.

Figure 4.14

Interaction Plot Comparing Change in Self-confidence Subscale



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. Pre-Post values indicate scores on the pre-survey (0) and post-survey (1). (Appendix F, p. 217)

The ANCOVA indicated no statistically significant difference between the groups. Since the residuals were normally distributed, the ANCOVA does have adequate power to conclude statistical significance between the Treatment and Control groups. Paired t-tests comparing pre-survey and post-survey responses were administered to confirm the results of the ANCOVA. (Table 4.24)

Table 4.24

Paired T-test Results for the Self-confidence Subscale Pre-test and Post-test

	N	Mean	Std Dev	Std Err	Minimum	Maximum	p-value
Control	36	-0.05	0.40	0.07	-0.57	1.50	0.44
Treatment	36	0.09	0.52	0.09	-1.93	1.29	0.33

Note: The negative minimum values indicate that some students rated the question higher on the pre-survey than on the post-survey. (Appendix F, p. 218)

The results are reflected in Table 4.24 and indicate that students in the Treatment group had a marginally higher score (M = 0.09, SD = 0.52) than students in the Control group (M = -0.05, SD = 0.40). Despite the normal distribution of the residuals and no statistical significance between the Treatment and Control groups' responses were established on the pre-survey, a two-sample t-test comparing post-survey responses on the Self-confidence subscale was administered to confirm the results of the ANCOVA. (Table 4.25)

Table 4.25

Post-survey Self-confidence Subscale Results

Group	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	36	3.06	0.32	0.05	2.36	3.93
Treatment	36	3.05	0.37	0.06	2.14	4.00

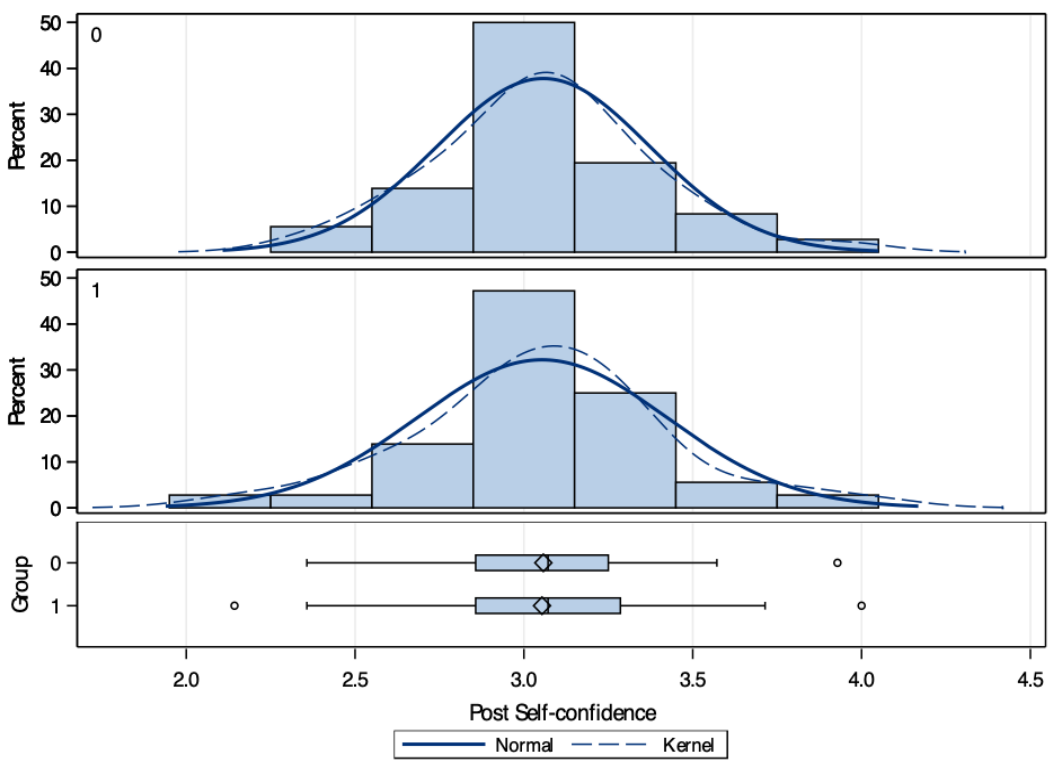
Note: Pooled $p = 0.96 > 0.05$ Satterthwaite $p = 0.96 > 0.05$ (Appendix F, p. 224)

The results are reflected in Table 4.25 and indicate that students in the Treatment group had a marginally lower score (M = 3.05, SD = 0.37) than students in the Control group

(M = 3.06, SD = 0.32), p = 0.96. The data was reasonably normally distributed as indicated in Figure 4.15. The box plots below the histogram indicate that the two groups significantly overlap with the means of both groups nearly identical.

Figure 4.15

Data Distribution for Post-survey Self-confidence Subscale



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. (Appendix F, p. 225)

Because the data did appear to be normally distributed, the parametric t-test has adequate power and a non-parametric Wilcoxon Rank Sums test was not necessary. The ANCOVA, combined with the two-sample t-test on the post-survey responses confirms that the null hypothesis fails to be rejected and there is no statistically significant difference between Treatment and Control groups on the post-survey Self-confidence subscale.

The subscale evaluating students' perceived value in learning to solve mathematics problems was evaluated using an ANCOVA which compared post-survey results to the students' responses to the same survey administered prior to the intervention. This test indicated the relative change in responses between the two groups. (Table 4.26)

Table 4.26

ANCOVA results for the Post-survey Value Subscale

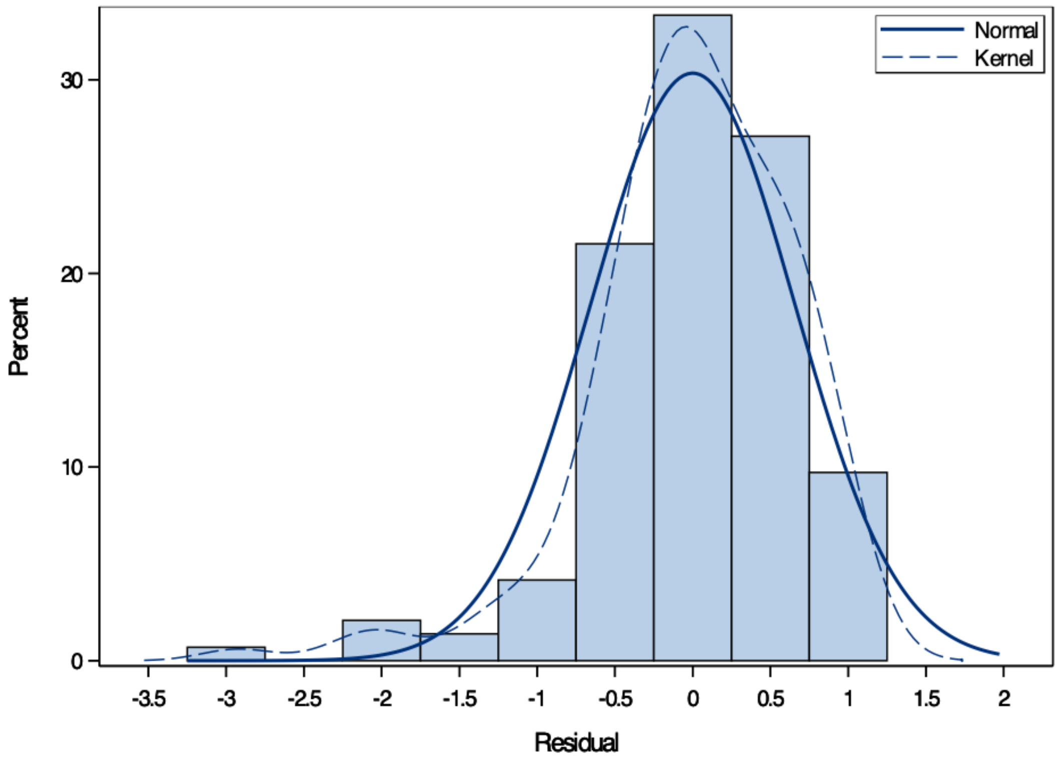
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.463	0.15	0.35	0.79
Error	140	61.81	0.44		
Corrected Total	143	62.27			

Note: The results of the ANCOVA indicate no statistically significant difference in survey responses between the Treatment and Control groups. (Appendix G, p. 230)

Table 4.26 reflects the results of the ANCOVA which show that experiencing the DAIP activity had no significant, positive effect on the students' Value subscale post-survey scores when controlling for survey Value subscale pre-survey scores, $F = 0.35$, $p < 0.79$. The ANCOVA analyzing the post-survey Value subscale indicates that the residuals do not have a normal distribution, so the parametric ANCOVA does not have adequate power to confirm significance. (Fig 4.16)

Figure 4.16

Distribution of Residuals the Post-survey Value Subscale

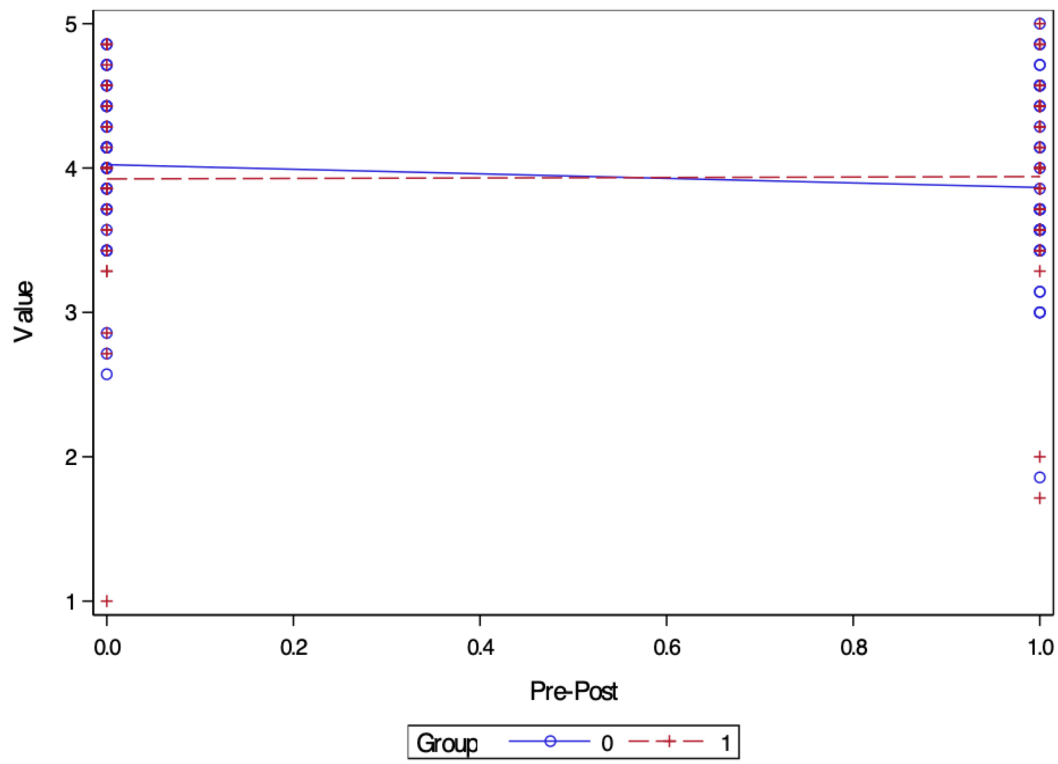


Note: (Appendix G, p. 234)

The Interaction Plot for the Value Subscale (Fig 4.17) compares the change in the two groups and indicates that following the intervention,indicates that following the intervention, both Control and Treatment groups students perceived similar value in learning to solve mathematics problems.

Figure 4.17

Interaction Plot Comparing Change in Value Subscale



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. Pre-Post values indicate scores on the pre-survey (0) and post-survey (1). (Appendix G, p. 235)

Despite a slight decrease in perceived value by the Control group and a steady perceived value by the Treatment group, the ANCOVA indicated no statistically significant difference between the groups. Since the residuals were not normally distributed, the ANCOVA does not have adequate power to confirm a statistical significance between the Treatment and Control groups. Paired t-tests comparing pre-survey and post-survey responses were run to confirm the results of the ANCOVA. (Table 4.27)

Table 4.27

Paired T-test Results for the Value Subscale Pre-test and Post-test

	N	Mean	Std Dev	Std Err	Minimum	Maximum	p-value
Control	36	-0.16	0.63	0.10	-1.86	1.71	0.14
Treatment	36	0.02	0.93	0.15	-3.14	2.57	0.92

Note: The negative minimum values indicate that some students rated the question higher on the pre-survey than on the post-survey. (Appendix G, p. 242)

The results are reflected in Table 4.27 and indicate that students in the Treatment group had a marginally higher score (M = 0.02, SD = 0.93) than students in the Control group (M = -0.16, SD = 0.63). No statistical significance between the Treatment and Control groups was established on the pre-survey responses. Since the residuals on the ANCOVA did not have a normal distribution, a two-sample t-test comparing post-test scores was run to confirm the results of the ANCOVA. (Table 4.28)

Table 4.28

Post-survey Value Subscale Results

Group	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	36	3.87	0.68	0.11	1.86	5.00
Treatment	36	3.94	0.67	0.11	1.71	5.00

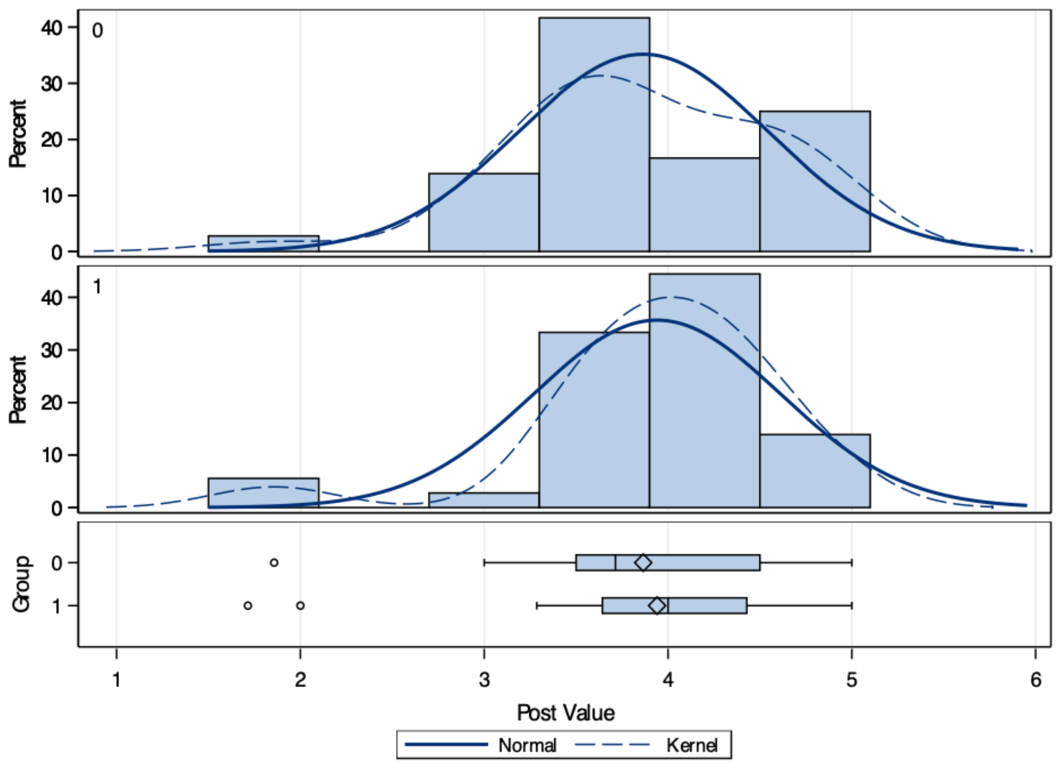
Note: Pooled $p = 0.64 > 0.05$ Satterthwaite $p = 0.64 > 0.05$ (Appendix G, p. 243)

The results are reflected in Table 4.28 and indicate that students in the Treatment group had a marginally higher score (M = 3.94, SD = 0.67) than students in the Control group

(M = 3.87, SD = 0.68), p = 0.64. The data was not normally distributed as indicated in Figure 4.18. The box plots below the histogram indicate that the two groups significantly overlap with the mean of the Treatment group slightly higher than that of the Control group.

Figure 4.18

Data Distribution for Post-survey Value Subscale



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. (Appendix G, p. 244)

Because the data did not appear to be normally distributed, a parametric test does not have adequate power to accurately predict significance and a non-parametric Wilcoxon Rank Sums test was also administered. (Table 4.29)

Table 4.29

Wilcoxon Rank Sums for Post-survey Value Subscale

Group	N	Sum of Scores	Expected Under H_0	Std Dev Under H_0	Mean Score
Control	36	1242.50	1314.0	88.48	34.51
Treatment	36	1385.50	1314.0	88.48	38.49

Note: $p = 0.43 > 0.05$ (Appendix G, p. 245)

The results for the Wilcoxon Rank Sums test reflected in Table 4.29 indicate that the Sum of Scores for the Treatment group (SoS=1385.50, Expected = 1314.0) is slightly higher than the Control group (SoS=1242.50, Expected=1314.0), $p = 0.43$. The results of the Wilcoxon test and the paired t-tests reaffirm the results of the parametric t-test and indicate that the null hypothesis fails to be rejected and conclude that there is no statistically significant difference between the responses of the Control and Treatment groups on the post-survey Value subscale.

The subscale evaluating students' enjoyment in learning to solve mathematics problems was evaluated using an ANCOVA which compared post-survey results to the students' responses to the same survey administered prior to the intervention. This test indicated the relative change in scores between the two groups. (Table 4.30)

Table 4.30

ANCOVA results for the Post-survey Enjoyment Subscale

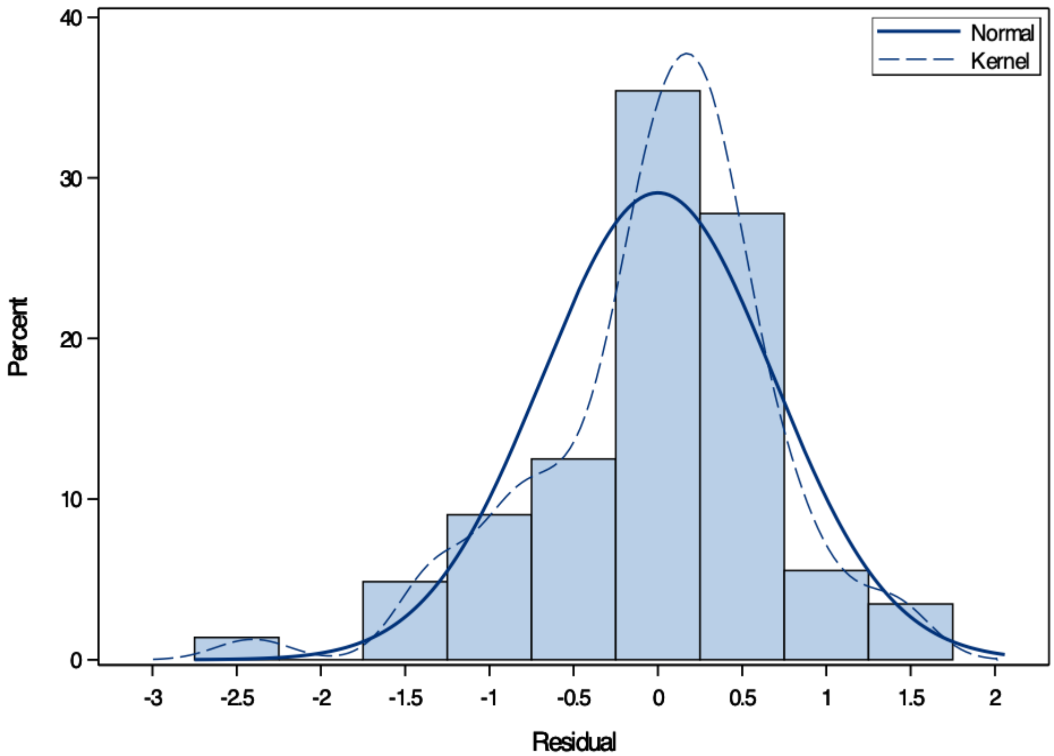
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.68	0.234	0.47	0.71
Error	140	67.348	0.48		
Corrected Total	143	68.018			

Note: The results of the ANCOVA indicate no statistically significant difference in survey responses between the Treatment and Control groups. (Appendix H, p. 249)

Table 4.30 reflects the results of the ANCOVA which show that experiencing the DAIP activity had no significant, positive effect on the students' Enjoyment subscale post-survey scores when controlling for survey Enjoyment subscale pre-survey scores, $F = 0.47$, $p < 0.71$. The histogram (Fig 4.19) shows that the residuals on the ANCOVA analyzing the post-survey Enjoyment subscale were not normally distributed.

Figure 4.19

Distribution of Residuals the Post-survey Enjoyment Subscale

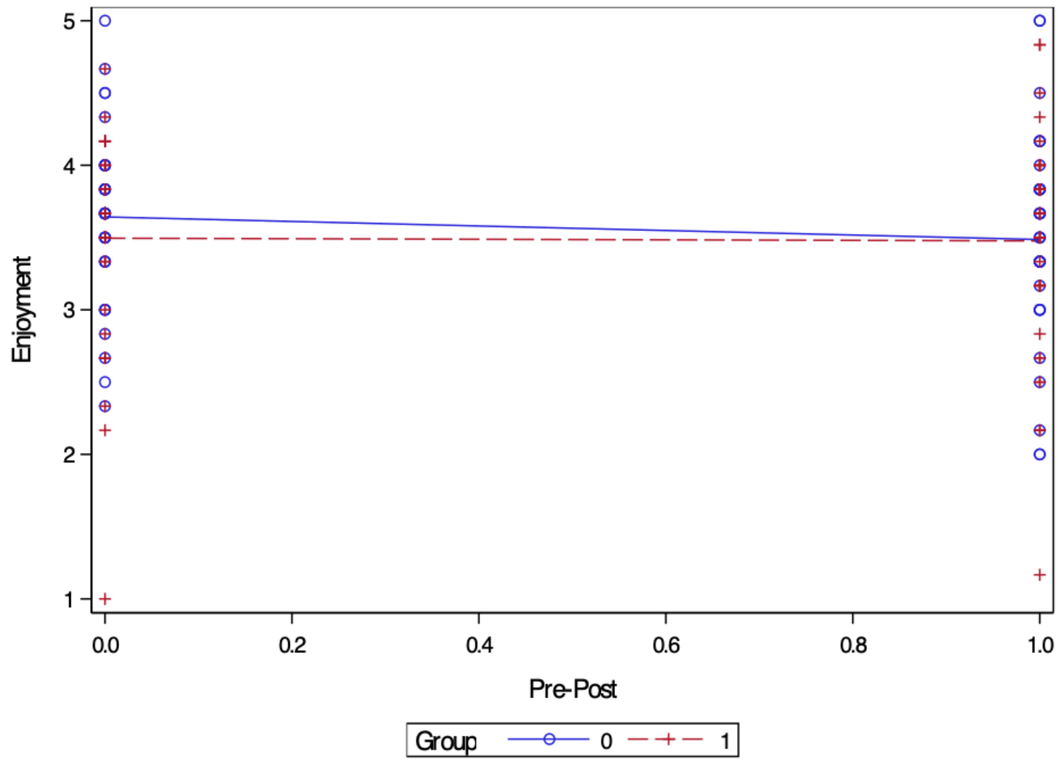


Note: (Appendix H, p. 253)

Despite the residuals not being normally distributed on the Enjoyment subscale, the Interaction Plot for the Enjoyment Subscale (Fig 4.20) compares the change in the two groups and indicates that following the intervention, both Control and Treatment groups experienced similar levels of enjoyment in solving mathematics problems, although the Treatment groups' self-confidence appears to have remained consistent while the Control groups' enjoyment declined.

Figure 4.20

Interaction Plot Comparing Change in Enjoyment Subscale



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. Pre-Post values indicate scores on the pre-survey (0) and post-survey (1). (Appendix H, p. 254)

Since the residuals were not normally distributed, the ANCOVA does not have adequate power to confirm a statistical significance between the Treatment and Control groups. Paired t-tests comparing pre-survey and post-survey responses were run to confirm the results of the ANCOVA. (Table 4.31)

Table 4.31

Paired T-test Results for the Enjoyment Subscale Pre-test and Post-test

	N	Mean	Std Dev	Std Err	Minimum	Maximum	p-value
Control	36	-0.16	0.60	0.10	-1.67	1.17	0.13
Treatment	36	-0.02	0.73	0.12	-1.50	3.00	0.13

Note: The negative minimum values indicate that some students rated the question higher on the pre-survey than on the post-survey. (Appendix H, p. 256)

The results are reflected in Table 4.31 and indicate that students in the Treatment group had a marginally higher score (M = -0.02, SD = 0.73) than students in the Control group (M = -0.16, SD = 0.60). Because no statistical significance between the Treatment and Control groups was established on the pre-survey responses and the residuals on the ANCOVA did not have a normal distribution, a two-sample t-test comparing post-survey responses was administered to confirm the results of the ANCOVA. (Table 4.32)

Table 4.32

Post-survey Enjoyment Subscale Results

Group	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	36	3.49	0.70	0.12	2.00	5.00
Treatment	36	3.48	0.77	0.13	1.17	4.83

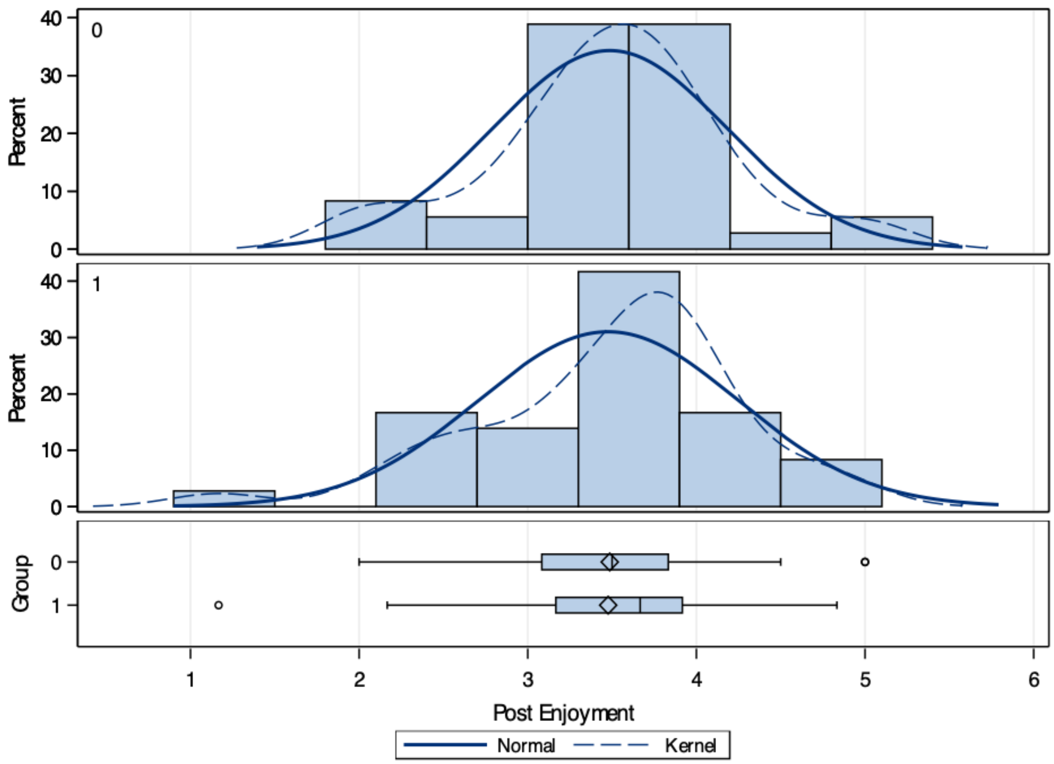
Note: Pooled $p = 0.96 > 0.05$ Satterthwaite $p = 0.96 > 0.05$ (Appendix H, p. 262)

The results are reflected in Table 4.31 and indicate that students in the Treatment group had a marginally higher score (M = 3.48, SD = 0.77) than students in the Control group

(M = 3.49, SD = 0.70), p = 0.96. The data was not normally distributed as indicated in Figure 4.21. The box plots below the histogram indicate that the two groups significantly overlap with the mean of the Treatment group slightly higher than that of the Control group.

Figure 4.21

Data Distribution for Post-survey Enjoyment Subscale



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. (Appendix H, p. 263)

Because the data did not appear to be normally distributed, a parametric t-test does not have adequate power to accurately predict significance and a non-parametric Wilcoxon Rank Sums test was also administered. (Table 4.33)

Table 4.33

Wilcoxon Rank Sums for Post-survey Enjoyment Subscale

Group	N	Sum of Scores	Expected Under H_0	Std Dev Under H_0	Mean Score
Control	36	1281.0	1314.0	88.39	35.58
Treatment	36	1347.0	1314.0	88.39	37.42

Note: $p = 0.71 > 0.05$ (Appendix H, p. 264)

The results for the Wilcoxon Rank Sums test reflected in Table 4.33 indicate that the Sum of Scores for the Treatment group (SoS=1347.0, Expected = 1314.0) is slightly higher than the Control group (SoS=1281.0, Expected=1314.0), $p = 0.71$. The results of the Wilcoxon test and the paired t-tests reaffirm the results of the parametric t-test and indicates that the null hypothesis fails to be rejected and conclude that there is no statistically significant difference between the responses of the Control and Treatment groups on the post-survey Enjoyment subscale.

The subscale evaluating students' motivation in learning to solve mathematics problems was evaluated using an ANCOVA which compared post-survey results to the students' responses to the same survey administered prior to the intervention. This test indicated the relative change in scores between the two groups. (Table 4.34)

Table 4.34

ANCOVA results for the Post-survey Motivation Subscale

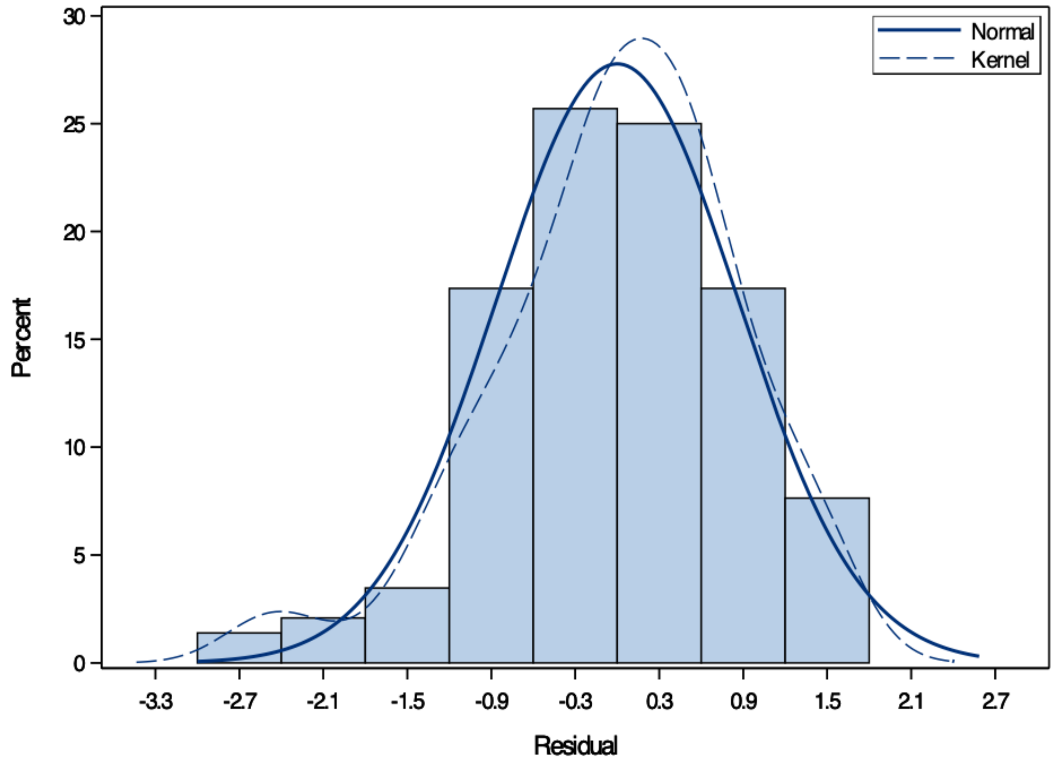
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1.62	0.54	0.71	0.55
Error	140	106.24	0.76		
Corrected Total	143	107.86			

Note: The results of the ANCOVA indicate no statistically significant difference in survey responses between the Treatment and Control groups. (Appendix J, p. 268)

Table 4.34 reflects the results of the ANCOVA which show that experiencing the DAIP activity had no significant, positive effect on the students' Motivation subscale post-survey scores when controlling for survey Motivation subscale pre-survey scores, $F = 0.71$, $p < 0.55$. The histogram (Fig 4.22) indicates that the residuals for the Motivation subscale were not normally distributed.

Figure 4.22

Distribution of Residuals the Post-survey Motivation Subscale

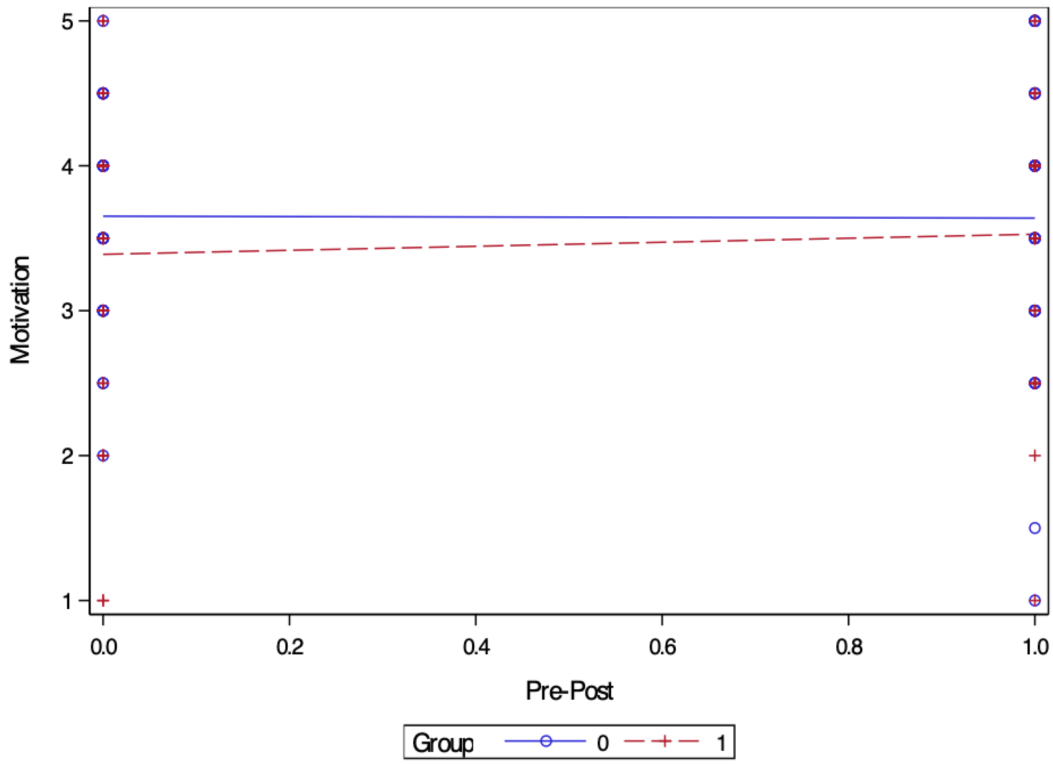


Note: (Appendix J, p. 272)

The Interaction Plot for the Motivation Subscale (Fig 4.23) compares the change in the two groups and indicates that following the intervention, both Control and Treatment groups experienced similar levels of motivation to solve mathematics problems, although the Treatment groups' motivation appears to have increased while the Control groups' self-confidence remained the same or slightly decreased.

Figure 4.23

Interaction Plot Comparing Change in Motivation Subscale



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. Pre-Post values indicate scores on the pre-survey (0) and post-survey (1). (Appendix J, p. 273)

Since the residuals were not normally distributed, the ANCOVA does not have adequate power to confirm a statistical significance between the Treatment and Control groups. Paired t-tests comparing pre-survey and post-survey responses were administered to confirm the results of the ANCOVA. (Table 4.35)

Table 4.35

Paired T-test Results for the Motivation Subscale Pre-test and Post-test

	N	Mean	Std Dev	Std Err	Minimum	Maximum	p-value
Control	36	-0.01	0.81	0.13	-2.00	2.50	0.92
Treatment	36	0.14	0.98	0.16	-3.00	2.50	0.40

Note: The negative minimum values indicate that some students rated the question higher on the pre-survey than on the post-survey. (Appendix J, p. 274)

The results are reflected in Table 4.35 and indicate that students in the Treatment group had a marginally higher score (M = 0.14, SD = 0.98) than students in the Control group (M = -0.01, SD = 0.81). No statistical significance between the Treatment and Control groups was established on the pre-survey responses, so a two-sample t-test comparing post-survey responses on the Motivation subscale was administered to confirm the results of the ANCOVA. (Table 4.36)

Table 4.36

Post-survey Motivation Subscale Results

Group	N	Mean	Std Dev	Std Err	Minimum	Maximum
Control	36	3.64	0.99	0.17	1.00	5.00
Treatment	36	3.53	0.84	0.14	1.00	5.00

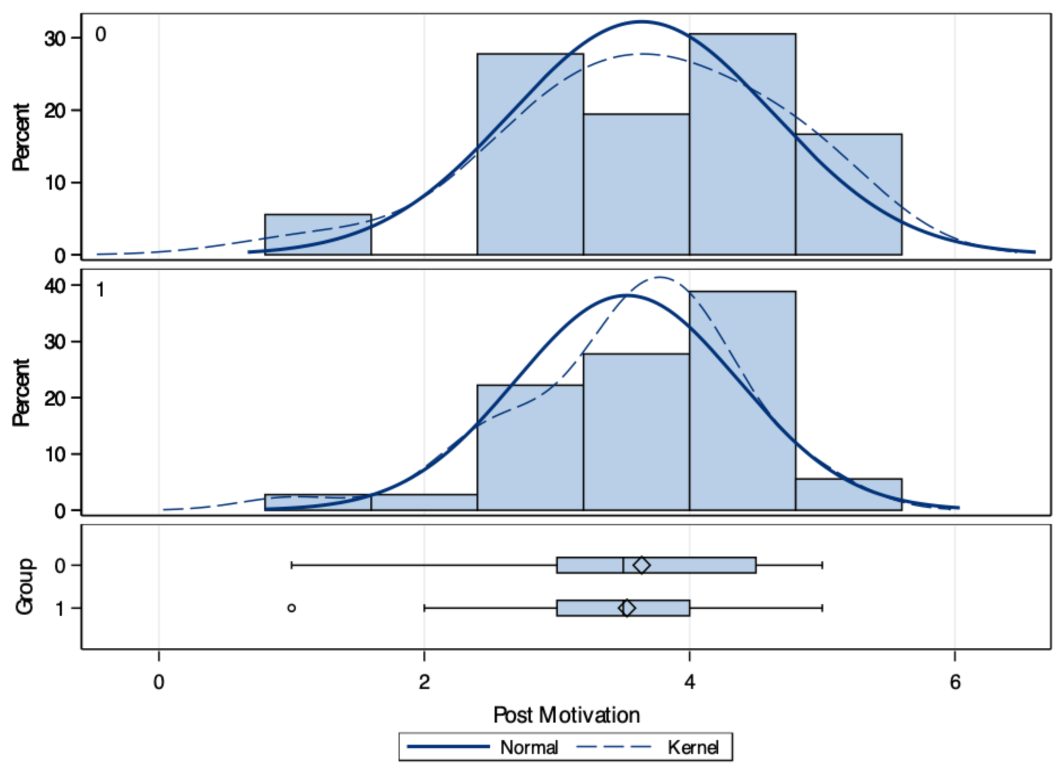
Note: Pooled $p = 0.61 > 0.05$ Satterthwaite $p = 0.61 > 0.05$ (Appendix J, p. 280)

The results are reflected in Table 4.36 and indicate that students in the Treatment group had a marginally higher score (M = 3.53, SD = 0.84) than students in the Control group

(M = 3.64, SD = 0.99), p = 0.61. The distribution of the data was not normally distributed as indicated in Figure 4.24. The box plots below the histogram indicate that the two groups significantly overlap with the mean of the Treatment group slightly lower than that of the Control group. Because the data did not appear to be normally distributed, a parametric test does not have adequate power to accurately predict significance and a non-parametric Wilcoxon Rank Sums test was also administered. (Table 4.37)

Figure 4.24

Data Distribution for Post-survey Motivation Subscale



Note: Group 0 represents the Control group, Group 1 represents the Treatment group. (Appendix J, p. 281)

Table 4.37

Wilcoxon Rank Sums for Post-survey Motivation Subscale

Group	N	Sum of Scores	Expected Under H_0	Std Dev Under H_0	Mean Score
Control	36	1360.0	1314.0	87.23	37.78
Treatment	36	1268.0	1314.0	87.23	35.22

Note: $p = 0.60 > 0.05$ (Appendix J, p. 282)

The results for the Wilcoxon Rank Sums test reflected in Table 4.37 indicate that the Sum of Scores for the Treatment group (SoS=1268.0, Expected = 1314.0) is slightly higher than the Control group (SoS=1360.0, Expected=1314.0), $p = 0.60$. The results of the Wilcoxon test reaffirm the results of the ANCOVA and the parametric t-test, indicating that the null hypothesis fails to be rejected and conclude that there is no statistically significant difference between the responses of the Control and Treatment groups on the post-survey Motivation subscale.

While the comparison between the groups' responses Attitudes Towards Mathematics Survey (ATMI) before and after the implementation of the DAIP activity do not indicate a statistically significant change, the results suggest a trend where those in the Treatment group experienced some improvement in confidence and interest while those in the Control group saw no improvement or slight decline in confidence and interest toward learning mathematics. (Table 4.38)

Table 4.38

ATMI Survey Confidence and Interest Summary

Pre-survey Result Summary

	Self-confidence		Value		Enjoyment		Motivation	
Group	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Control	3.11	0.40	4.02	0.56	3.64	0.58	3.65	0.66
Treatment	2.97	0.37	3.92	0.73	3.50	0.71	3.39	0.96
p-value	p = 0.13> 0.05		p = 0.52> 0.05		p = 0.34> 0.05		p = 0.18> 0.05	

Note: N=36 for each group, p-values indicate no statistical difference between groups for any subscale.

Post-survey Result Summary

	Self-confidence		Value		Enjoyment		Motivation	
Group	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Control	3.06	0.32	3.87	0.68	3.49	0.70	3.64	0.99
Treatment	3.05	0.37	3.94	0.67	3.48	0.77	3.53	0.84
p-value	p = 0.96> 0.05		p = 0.64> 0.05		p = 0.96> 0.05		p = 0.60> 0.05	

Note: N=36 for each group, p-values indicate no statistical difference between groups for any subscale.

Summary

While the Treatment group scored marginally higher than the Control group on the Renaissance STAR exam and the teacher-generated assessment after the implementation of the DAIP activity, results indicate that there is no statistically significant difference between the groups' performance on either assessment. Similarly, subjects' responses on the Attitudes Towards Mathematics Survey (ATMI) before and after the implementation of the DAIP activity do not indicate a statistically significant change. Although not statistically significant, the results suggest a trend where those in the Treatment group experienced some improvement in confidence and interest while those in the Control group saw no improvement or slight decline in confidence and interest towards learning mathematics.

Chapter 5
Discussion

Introduction

The underperformance of African American students in underprivileged schools is well-documented (Darling-Hammond, 2000). Inequitable access to resources limits the success of some African American students (Anderson, 2017). Barriers continue to curtail access to equal education opportunities for this demographic in particular (Allen and Jewell, 1995). Many urban and suburban school systems have a disparity in academic resources available to their respective staff and students due to the community in which they reside (Jenkins, 2017). Since these underprivileged schools will continue to struggle to match physical and academic resources available to their suburban counterparts, alternative and low-cost strategies must be explored to offset the disparities and provide opportunities for African American students to engage and succeed in STEM fields.

Design Arts Integrated Project (DAIP) is a teaching strategy that shows potential as one such alternative and low-cost strategy. As one component of DAIP, project-based learning, students exercise cross-disciplinary skills through the development of a solution to a real-world problem. Solved with design inquiry, targeted learning experiences develop stronger critical thinking skills. The integration of arts components to express learning allows the students to utilize non-traditional means of articulating their learning. While anecdotal evidence exists pointing to the effectiveness of PBL in engaging students (Allen et. al 2013), little research had been done quantifying the effectiveness with the middle school demographic and virtually none with African American students. By testing the use of DAIP which incorporates PBL, design thinking, and expressed through arts to teach mathematics in a middle school pre-algebra classroom in a 95-percent African American district, the efficacy of this alternative teaching strategy was examined.

One of the introductory components of the project-based learning activity consisted of students designing a function machine using an array of art supplies. During the introduction to the unit, the teacher demonstrated a model of a function machine that was hand-crafted from a cereal box. The concept of a function machine fascinated the students. The students then collaborated in pairs to determine the function rule as one value was placed in the machine and another value came out. The teacher provided minimal assistance as the students observed the model and worked to build their machines and create input/output values with increasingly complex rules for their peers to ultimately decipher. As the DAIP evolved, students came to discover that they were creating and solving functions. While designing and testing their machines, the students demonstrated a higher level of engagement than their counterparts in the control group which was concurrently receiving the same material employing a traditional teaching strategy.

Following the function machine lesson, the students in the Treatment group were introduced to the second milestone in the DAIP; participating in a Math-A-Thon that supported a national children's research hospital. Developing the plan to participate in the Math-A-Thon firmly planted the concept of functions in their minds. Using their autonomous creativity to individually solicit sponsors and donors to raise money, they worked through solving problems with varying x-values to reach an intended goal of their own choosing. Some students welcomed this freedom offered by the design process, while others sought constant guidance and approval. Students were required to create the media by which they solicited pledges and donations. This creative freedom resonated with many of the students, deepening their commitment to the project and resulted in artifacts that varied from videos to slideshow presentations to posters and informational brochures.

Students were regularly challenged to think critically when presented with tables, graphs, and equations to interpret, analyze, and solve for missing values. As the DAIP continued, students who had secured their donors and sponsors were arranged into groups of three to four members and charged with graphing, displaying, and presenting both their individual and groups' pledge collections. They had to collaborate to determine a function rule encompassing the total amount of money the group solicited, an exercise that was quite challenging. The students seemed comfortable in understanding their individual functions as it related to their fundraising efforts; however, the students were initially confused when asked to pool the pledges as a group and to work in reverse to determine the potential x-value for each Math-A-Thon problem completed or attempted. In many instances, a leader emerged in a group and was able to guide their teammates. Other groups required a little coaching by the teacher. Students experienced discovery learning and constructivism firsthand as they were given a set of parameters to adhere to and execute without a clear procedure defined.

Summary of Findings

Research Question 1: To what extent would minority students' STEM scores in a disenfranchised community improve when taught STEM concepts using an alternative problem-solving teaching methodology (DAIP)?

The initial findings do not point to a clear advantage of one teaching strategy over the other. Despite no statistically significant difference between the Treatment and Control groups' scores on a standardized test or on a teacher-generated assessment, both measures reflected that the students in the Treatment group who participated in the project-based learning activity did perform better on both measures, scoring 2.08% higher than the Control group on the standardized test and 2.17% higher than the Control group on the teacher-generated assessment.
(Table 5.1) According to the classroom teacher implementing the lessons, the transient nature of the particular sample population studied has historically returned test results that varied widely in large part because of inconsistent daily attendance. A logical conclusion can be drawn that this pattern would lead to the data not being normally distributed. Another factor that may have contributed to less difference in student performance could be discomfort with the alternative strategy. Discovery learning and design thinking require the student to embrace making mistakes, something penalized in empirical instruction and difficult for them to accept.

Table 5.1

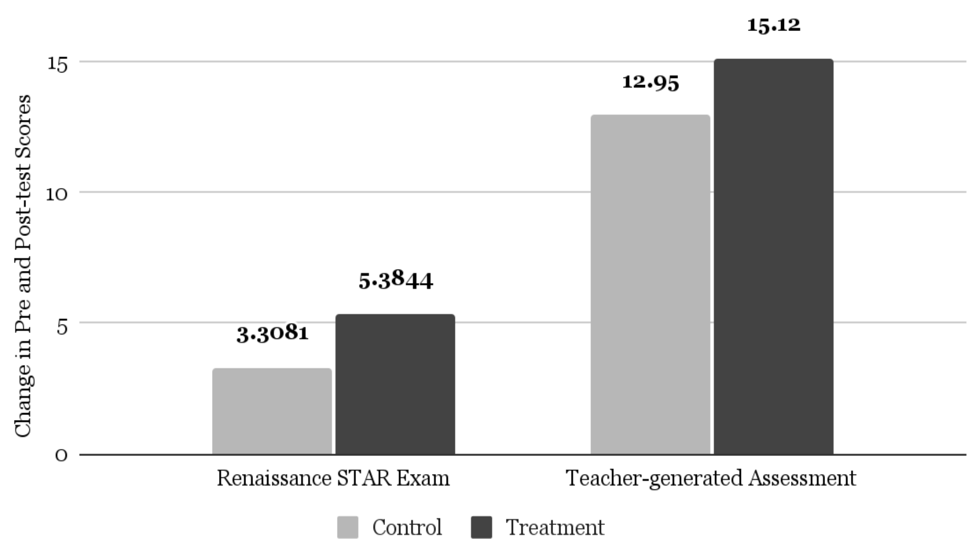
Change in Score Summary on the Renaissance STAR and Teacher-generated Assessments

	Renaissance STAR Exam			**Teacher-generated Assessment**		
Group	Pre-Mean	Post-Mean	Change	Pre-Mean	Post-Mean	Change
Control	13.44	16.75	3.31	22.86	35.81	12.95
Treatment	15.15	20.54	5.38	26.1	41.22	15.12

A relatively small sample size of N=71 on the standardized test and N=72 on the teacher-generated assessment combined with the research being conducted on a single, eight-week unit of study were factors that likely contributed to seeing no statistically significant difference between groups. The slight increase in performance by the students completing the DAIP seems to indicate that the alternative strategy shows promise and warrants additional research. (Figure 5.1) A larger sample size combined with a longer duration may provide more regularly distributed data and a more significant difference in test scores.

Figure 5.1

Change in Means on the Renaissance STAR and Teacher-generated Assessments



Research Question 2: To what degree would minority students in a disenfranchised community improve in confidence and interest in STEM concepts when taught mathematics using an alternative problem-solving strategy (DAIP)?

Findings indicate no statistically significant difference in confidence and interest in STEM concepts when being taught mathematics using DAIP instead of a traditional teaching strategy. Where subjects in the Treatment group showed specifically higher engagement was in using the arts and the making of media to demonstrate their learning. While higher enthusiasm and engagement among the students in the Treatment group was observed during the intervention, the responses given on the post-survey were only slightly higher than those of the students in the Control group. (Table 5.2)

Table 5.2

Change in Response Summary on the Attitudes Toward Mathematics Inventory Survey

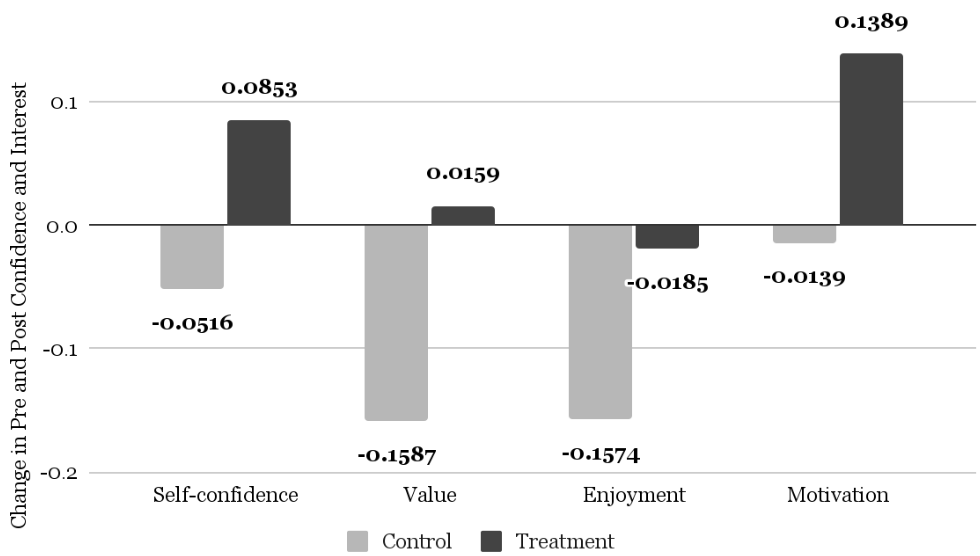
	Self-confidence			**Value**		
Group	Pre- Mean	Post- Mean	Change	Pre- Mean	Post- Mean	Change
Control	3.11	3.06	-0.05	4.02	3.87	-0.16
Treatment	2.97	3.05	0.09	3.92	3.94	0.02

	Enjoyment			**Motivation**		
Group	Pre- Mean	Post- Mean	Change	Pre- Mean	Post- Mean	Change
Control	3.64	3.49	-0.16	3.65	3.64	-0.01
Treatment	3.50	3.48	-0.02	3.39	3.53	0.14

While not statistically significant, the changes on the subscales Self-confidence, Value, and Motivation all increased for the Treatment group and decreased for the Control group. The Control group indicated that their confidence and interest were lower after learning about functions than before while the Treatment group responded that their Self-confidence, Value, and Motivation had increased. Both groups responded that their Enjoyment declined, although the decline in the Treatment group was less than the decline in Enjoyment experienced by the Control group. (Figure 5.2) In observing the Treatment groups' participation in the DAIP, the decrease in Enjoyment could be attributed to the overall feeling of frustration many of the students expressed as they were not accustomed to the level of rigor this type of instruction required. Some students did not see the benefit in constantly adjusting their responses to either match a given rule or determine a rule based on input values.

Figure 5.2

Change in Means on the Attitudes Toward Mathematics Inventory Survey



The same factors at play with the scores on the Renaissance STAR and teacher-generated assessments appear to influence the students' responses on the survey. The transient nature of the particular sample population studied has historically returned test results that varied widely in large part because of inconsistent daily attendance. A logical conclusion can be drawn that this pattern would lead to widely varied survey responses. Another factor that may have contributed to less difference in student Enjoyment could be discomfort induced by the iterative design process used in the alternative strategy. Embracing mistakes, something penalized in explicit instruction, requires a significant paradigm shift in young students.

A relatively small sample size (N=72) combined with the research being conducted on a single, eight-week unit of study were factors that likely contributed to seeing no statistically significant difference between groups. The slight increase in confidence and interest in learning mathematics by the students completing the DAIP seems to indicate that the alternative strategy shows promise and warrants additional research. A larger sample size combined with a longer duration may provide more regularly distributed data and a more significant difference in test scores.

Integration of Findings with Current Research

While the differences between the Treatment and Control groups were not large enough to return statistically significant results, the findings seem to align with other research done on project-based learning. Previous research has shown that students involved in PBL tend to demonstrate a higher level of motivation than those individuals who do not engage in a PBL experience (Blumenfeld et al., 1991). Student responses on the ATMI survey, while only slightly higher than the students in the Control group, appear to reinforce the data found in the literature.

Successful PBL requires a commitment to the real-world topic and is necessary for the PBL to be authentic. By collaborating on decisions of the project from start to finish, the students make their own assumptions on how to complete a project, vesting their interest in the results (Robinson, 2013). The students in the Treatment group demonstrated an observed vested interest in raising money for the children's hospital and daily engagement in the activity was observed to be higher than their counterparts in a section receiving traditional instruction.

A particularly potent portion of the DAIP was the requirement that the students used artistic expression to solicit pledges and demonstrate their fund-raising success. Research indicates that integrating arts into the learning bridges the gaps between abstract and concrete conceptual ideas (Marshall, 2014). Arts integration can have a particularly positive impact on disenfranchised learners (Duma et. al, 2014) and in this research, the subjects found enjoyment in learning and expressing functions through art.

Little research has been conducted in the use of design thinking with students younger than at the collegiate level. Historically employed in the business, engineering, and design curricula of professional programs, design thinking as a problem-solving process has little documented use with young students when studying traditional subjects. Many students required coaching and encouragement, reassuring them that making mistakes was not only acceptable, but expected and part of the problem-solving process. Nonetheless, this research demonstrates that students at a younger age can successfully use the principles of design thinking to solve complex problems.

Implications for Practice

Based on the experience in executing the intervention, the findings suggest the following implications:

1. Effectively implementing project-based learning shows potential as a catalyst to motivate and improve performance outcomes within a disenfranchised student population. Its cost-effectiveness lends itself to sustainability, especially in schools where resources are limited.
2. Educators who are unfamiliar with the numerous components of the DAIP teaching methodology will need to seek professional development opportunities to successfully guide students through this process. Some areas requiring additional training or reinforcement include; crafting a meaningful DAIP from a

real-world problem, facilitating the design process, and conditioning students to understand and accept the failures inherent in a trial-and-error method of inquiry.

3. The iterative activity of design thinking in conjunction with project-based learning can lead to frustration for students. Most students and teachers are not accustomed to experiencing the expected failures of the trial-and-error nature of the process. Participants must be conditioned to these revised expectations.
4. To be able to grow confidence and interest in mathematics, the students need to experience success out of the failures encountered in the process. While the students experience discovery learning, the learning process must be carefully crafted and guided by the teacher for the learning to be authentic.

Recommendations for Future Research

Despite the inconclusive results of this particular research project, a trend can be observed when looking at all six measures as a pattern. This pattern seems to indicate that the use of design thinking to solve a real-world problem and expressing learning through art holds promise. A few key limitations such as a modest sample size and short duration of the intervention likely contributed to results that were not statistically significant. One opportunity for future research would be to expand the sample size. A larger sample pool would likely return more regularly distributed data and may provide more conclusive results. Additionally, by repeating the design process to solve problems multiple times, the students become more comfortable and familiar with the technique. Extending the use of DAIP through multiple teaching units, perhaps for a semester or even a year may provide a more sustained and definitive pattern.

Since the DAIP teaching strategy employed does show promise, another area to be explored is whether similar results can be achieved with a different demographic. Value could be gained by determining if suburban students or rural students, or if students of other ages might benefit equally from this technique. Finally, expansion of the research may entail applying the teaching strategy to other subject matter. The cross-disciplinary nature of DAIP implies that the strategy could be successful in developing knowledge in other fields and even teach multiple concepts and skills in several disciplines simultaneously.

Conclusion

In an academic environment of ever-tightening budgets and limited resources, educators continue to search for innovative ways to effectively build knowledge and perhaps, more importantly, critical-thinking and problem-solving skills in learners. Teaching strategies that foster learning through the exercise of creativity and inquiry have the potential of efficiently building that knowledge while simultaneously developing needed skills. Research indicates that attaching learning to a real-world problem that is relevant to the student is more likely to engage that student (Cervantes et al., 2015). Well-conceived DAIP activities can fulfill these tasks with minimal cost.

The iterative nature of design thinking as a problem-solving technique gives students a unique ability to tackle problems that contain multiple variables with multiple solutions. By enhancing the creativity employed in this process with an arts component, students are engaged in non-linear inquiry and discovery that has the potential to develop advanced skills that can be applied to lifelong learning, long after their academic careers have ended.

The goal was to provide educators with another effective tool that can connect students in underserved areas with STEM topics. Allowing students to explore a problem that has real meaning in their lives and exercise creativity in finding a solution unlocks a belief in African American learners that they are capable of succeeding in areas where they have historically been underrepresented.

References

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, 103(1), 1–18. https://doi.org/10.1037/a0021017.

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system—secondary. SchoolPsychology Review, 42(1), 76-88.

Allen, W. O., & Jewell, J. O. (1995). American Education Since "An American Dilemma." *Daedalus*, *124*(1), 77–100.

Anderson, C. (2017). *White rage: The unspoken truth of our racial divide*. Bloomsbury, an imprint of Bloomsbury Publishing Plc.

Baird, D. (2015). *Integrating the Arts in Mathematics Teaching*. University of Toronto.

Bell, S. (2010). Project-based learning for the 21st century: Skills for the future. Clearing House, 83(2), 39-43; doi:10.1080/00098650903505415

Blumenfeld, P. C., Soloway, E., Marx, R., W., Krajcik, J. S., Guzdial, M., Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. Educational Psychologist, 26(3), 369-398.

Brazee, E. (Ed.). (2000). *Exploratory Curriculum in the Middle School*. ERIC Clearinghouse on Elementary and Early Childhood Education Champagne, Il.

Bruner, Jerome S. (1960) *The Process of Education: A Landmark in Educational Theory*. Harvard University Press.

Bruner, Jerome S. (1961). The Act of Discovery. *Harvard Educational Review*, *31*, 21–32.

Bruner, Jerome S. (1996) *The Culture of Education*. Harvard University Press.

Buck Institute for Education. (2014). Resource list: Managing a project. Retrieved from http://www.bie.org/blog/resource_list_management_in_pbl

Capraro, R. M. & Slough, S. W. (2008). Why PBL? Why STEM? Why now? An introduction to project-based learning: An integrated science, technology, engineering, and mathematics (STEM) approach. In R. M. Capraro & S. W. Slough (Eds.), Project based learning: An integrated science technology engineering and mathematics (STEM) approach (pp. 1–6). Rotterdam: Sense

Carroll, M. P., Ph. D. (2014). Shoot for the moon! The mentors and the middle schoolers explore the intersection of design thinking and STEM. *Journal of Pre-College Engineering Education Research*, *4*(1).

Cervantes, B., Hemmer, L., & Kouzekanani, K. (2015). The Impact of Project-Based Learning on Minority Student Achievement: Implications for School Redesign. *E National Council of Professors of Educational Administration*, *2*(2), 50–66.

Chine, D. (2022). *A Pathway to Success? A Longitudinal Study Using Hierarchical Linear Modeling of Student and School Effects on Academic Achievement in a Middle School STEM Program*.

Clark, A. C. & Ernst, J. V. (2007). A model for the integration of science, technology, engineering and mathematics. The Technology Teacher, 66(4), 24–26.

Cody, Ron. (2021) *A Gentle Introduction to Statistics Using SAS Studio in the Cloud*. Cary, NC: SAS Institute Inc.

College Board. (2013). College bound seniors: Total group profile report. New York, NY: College Board.

Condliffe, B., Quint, J., Visher, M. G., Bangser, M. R., Drohojowska, S., Saco, L., & Nelson, E. (2017). *Project-Based Learning A Literature Review*.

Cooper, E. & Sherk, J. (1989). Addressing urban school reform: Issues and alliances. Journal of Negro Education, 58 (3), 315-331.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: Sage Publications, Inc.

Culata, R. (2018, November 30). Constructivist Theory (Jerome Bruner). InstructionalDesign.org. https://www.instructionaldesign.org/theories/constructivist/.

Dabney KP, Tai RH, Almarode JT, Miller-Friedmann JL, Sonnert G, Sadler PM, Hazari Z (2012) Out-of-school time science activities and their association with career interest in STEM. Int J Sci Educ Part B 2(l):63-79.

Dahl, G. B., & Lochner, L. (2012). *The impact of family income on child achievement*: Evidence from the Earned Income Tax Credit. *American Economic Review*, *102*(5), 1927–1956.

Dantley, S. J., & Leonard, J. (2010). The State of Mathematics and Science Achievement among African American Males. In E. M. Zamani-Gallaher & V. C. Polite (Eds.), *The State of the African American Male* (pp. 45–63). Michigan State University Press.

Darling-Hammond, L., Barron, B., Pearson, P. D., Schoenfeld, A. H., Stage, E. K., Zimmerman, T. D., Cervetti, G. N., and Tilson, J. L. (2008). Powerful learning: 67 *What we know about teaching for understanding*.San Francisco, CA: Jossey-Bass

Darling-Hammond, L. (2000). New standards and old inequalities: School reform and the education of African American students. *The Journal of Negro Education*, *69*(4), 263.

Darling-Hammond, L. (1997). The right to learn: A blueprint for creating schools that work. San Francisco: Jossey-Bass.

Dolmans, D., Grave, W. D., Wolfhagen, I. & Vleuten, C. (2005). Problem based learning: Future challenges for educational practice and research. Medical Education, 39, 732–741.

Dorst, Kees (2011). The core of 'design thinking' and its application. *Design Studies,* 32(6). (pp.521-532). doi:10.1016/j.destud.2011.07.006

Duch, B. J., Groh, S. E., & Allen, D. E. (2001). *Why problem-based learning? A case study of institutional change in undergraduate education*. In B. J. Duch, S. E.

Groh, & D. E. Allen (Eds.), The power of problem-based learning (pp. 3-1 1). Sterling: VA: Stylus.

Duma, A., & Silverstein, L. (2014). A view into a decade of arts integration. *Journal for Learning through the Arts A Research Journal on Arts Integration in Schools and Communities*, *10*(1). https://doi.org/10.21977/d910119197

Efland, A. (2002). *Art and cognition: Integrating the visual arts in the curriculum*. Teachers' College Press.

Egenrieder, J. (2010). Facilitating Student Autonomy in Project-Based Learning to Foster Interest and Resilience in STEM Education and STEM Careers. *Journal of the Washington Academy of Sciences*, *96*(4), 35–45.

Gonzalez, H. B., & Kuenzi, J. J. (2012, August). Science, technology, engineering, and mathematics (STEM) education: A primer. Washington, DC: Congressional Research Service, Library of Congress.

Grushka, K., & Curtis, N. (2018). Visual Art, Visual Design and Numeracy. In *Numeracy in Authentic Contexts: Making Meaning Across the Curriculum* (pp. 423–429). Singapore, Springer Nature.

Hanif, S., Wijaya, A. F. C., & Winarno, N. (2019). Enhancing students' creativity through STEM project-based learning. *Journal of Science Learning*, *2*(2), 50.

Hannemann A (2007) Incubating innovation: Diversity efforts rejuvenate life science work force. Science 5(11). http://sciencecareers.sciencemag.org/career_magazine/previous_issues/articles/2007_05_ 1 1 /science, opms .r0700033

Henriksen, D. (2017). *Creating STEAM with Design Thinking: Beyond STEM and Arts Integration.* Steam, 3(1), 1–11. https://doi.org/10.5642/steam.20170301.11

Hertzog, N. B. (2007). *Transporting pedagogy: Implementing the project approach in two first-grade classrooms.* Journal of Advanced Academics, 18(4), 530-564.

Hetland, L., & Winner, E. (2004). Cognitive transfer from arts education to non-arts outcomes: Research evidence and policy implications. In E. Eisner & M. Day (Eds.) Handbook of Research and Policy in art education (pp. 135-161). Mahwah, NJ: Elbaum and Associates.

Hmelo-Silver, C. E. (2004). *Problem-Based Learning: What and How Students Do Learn?* Educational Psychology Review, 16(3), 235–266.

Horan, C., Lavaroni, C. & Beldon, P. (1996). *Observation of the tinker tech program students for critical thinking and social participation behaviors.* Novato: Buck Institute for Education.

Jenkins, J. (2017). T*he Effectiveness of Project-Based Learning on Mathematics Proficiency with African American Students.* Mercer University. Jonassen, 2000.

Johnson, C., & Kritsonis, W. A. (2006). The National Dilemma Of African American Students: Disparities In Mathematics Achievement And Instruction. *National Forum of Applied Educational Research Journal*, *20*(3), 1–8.

Joshi, A., Kale, S., Chandel, S., & Pal, D. (2015). *Likert scale: Explored and explained.* British Journal of Applied Science & Technology, 7(4), 396–403. https://doi.org/10.9734/bjast/2015/14975

Kabilan, M. K., & Kamarrudin, F. (2010). Engaging learners' comprehension, interest and motivation to learn literature using the reader's theatre. *Teaching English: Practice and Critique* (9), 132-159.

Kimbell, Lucy. (2011). Rethinking Design Thinking: Part I. Design and Culture. 3. 285-306. 10.2752/175470811X13071166525216.

Krajcik, J. S., Blumenfeld, P. C., Marx, R.W., & Soloway, E. (1994). A collaborative model for helping middle grade science teachers learn project-based instruction. The Elementary School Journal 94(5), 483-497.

Ladson-Billings, G. (1994). What Can We Learn From Multicultural Education. *Educational Leadership*, *51*(8), 22–26.

Laerd Statistics. (2018). One-way ANCOVA in SPSS Statistics [website]. Retrieved from https://statistics.laerd.com/spss-tutorials/ancova-using-spss-statistics.php

Lathan, J. (n.d.). *Why STEAM is so Important to 21st Century Education*. University of San Diego. Retrieved April 10, 2021. https://onlinedegrees.sandiego.edu/steam-education-in-schools/#STEAM

Lau, K. J. (2014). *The Effect of Art-Integrated Math Lessons on Student Interest and Engagement*. Saint Mary's College.

Lee, V. E., & Burkam, D. T. (2002). *Inequality at the starting gate: Social background* differences *in achievement as children begin school.* Retrieved from https://eric.ed.gov/?id=ED470551/

Leverenz, C. S. (2014). Design thinking and the wicked problem of teaching writing. Computers and Composition. https://doi.org/10.1016/j.compcom.2014.07.001

Li, Y., Schoenfeld, A. H., diSessa, A. A., Graesser, A. C., Benson, L. C., English, L. D., & Duschl, R. A. (2019). *Design and Design Thinking in STEM Education*. Journal for STEM Education Research, 2(2), 93–104. https://doi.org/10.1007/s41979-019-00020-z

Luka, I. (2020). Design Thinking in Pedagogy. *Journal of Education Culture and Society*, *5*(2), 63–74. https://doi.org/10.15503/jecs20142.63.74

Maarif, S. (2016). Improving junior high school students' mathematical analogical ability using discovery learning method. *International Journal of Research in Education and Science (IJRES)*, 2(1), 114- 124.

MaKinster, J. G., Barab, S. A., & Keating, T. M. (2001). Design and implementation of an on-line professional development community: A Project-Based Learning approach in a graduate seminar. *Electronic Journal of Science Education*, 5(3), 1-8.

Marshall, J. (2014). Transdisciplinarity and art integration: Toward a new understanding of art-based learning across the curriculum. *Studies in Art Education*, *55*(2), 104–127.

McGrath, D. (2004). Strengthening collaborative work. *Learning and Leading with Technology*, 31(5), 30-33.

McLeod, S. A. (2018, June 06). Jean piaget's theory of cognitive development. Simply Psychology. https://www.simplypsychology.org/piaget.html

McLeod, S. A. (2018, August 05). Lev Vygotsky. Simply Psychology. https://www.simplypsychology.org/vygotsky.html

Melles, G., Howard, Z., & Thompson-Whitesidec, S. (2012). Teaching design thinking: Expanding horizons in design education. Procedia - Social and Behavioral Sciences. https://doi.org/10.1016/j.sbspro.2011.12.035

Mergendoller, J. R., Maxwell, N. L., & Bellisimo, Y. (2006). The effectiveness of problem-based instruction: A comparative study of instructional methods and student characteristics. Interdisciplinary *Journal of Problem-based Learning*, 1(2), 49-69.

Mizell, H. (2010). *Why Professional Development Matters* (V. von Frank, Ed.). Learning Forward.

Mor, Y., & Mogilevsky, O. (2013). The learning design studio: collaborative design inquiry as teachers' professional development. *Research in Learning Technology*, *21*.

Moore, C., & Linder, S. (2012). Using dance to deepen student understanding of geometry. *Journal of Dance Education*, 12(3), 104-108.

Moreno, N. P., Tharp, B. Z., Vogt, G., Newell, A. D., & Burnett, C. A. (2016). Preparing Students for Middle School Through After-School STEM Activities. *Journal of Science Education and Technology*, *25*(6), 889–898.

Mosely, G., Wright, N., & Wrigley, C. (2018). Facilitating design thinking: A comparison of design expertise. Thinking Skills and Creativity. https://doi.org/10.1016/j.tsc.2018.02.004

Museus SD, Palmer RT, Davis RJ, Maramba D (2011) Racial and ethnic minority students' success in STEM education (ASHE Higher Education Report Vol. 36, No. 6). Jossey-Bass, Hoboken. doi : 10.1002/aehe. 3606

National Archives and Records Administration. (2022, February 8). *Plessy v. Ferguson (1896)*. National Archives and Records Administration. https://www.archives.gov/milestone-documents/plessy-v-ferguson?_ga=2.73818843.852258588.1645497381-935454520.1645497381

National Center for Educational Statistics. (2019, February). *Indicator 11: Mathematics achievement*. Status and Trends in the Education of Racial and Ethic Groups. https://nces.ed.gov/programs/raceindicators/indicator_rcb.asp

National Center on Intensive Intervention. (2018). STAR Math. STAR Math Performance Level Standards. https://charts.intensiveintervention.org/progressmonitoring/tool/?id=bd782b6fd24267c6.

Pande, M., & Bharathi, S. V. (2020). Theoretical foundations of design thinking – A constructivism learning approach to design thinking. Thinking Skills and Creativity, 36 (February), 100637. https://doi.org/10.1016/j.tsc.2020.100637

Piaget, J., & Cook, M. (1957). *The construction of reality in the child*. Basic Books.

Plattner, H. (ed.). (2010). D.school Bootcamp Bootleg. Institute of design at Stanford. Retrieved from: http://dschool.stanford.edu/wp-content/uploads/2011/03/BootcampBootleg2010v2SLIM.pdf

Plattner, H. Meinel, C., & Weinberg, U. (2009). Design Thinking. Innovation lernen – Ideenwelten öffnen. Mi-Wirtschaftsverlag.

Renaissance. (1998). *STAR Math*. Renaissance. https://www.renaissance.com/products/star-math/.

Riley, S. (2013, August 11). *K-12 professional development*. Artsintegration.Com. https://artsintegration.com/

Robinson, J. (2013). Project-based learning: Improving student engagement and performance in the laboratory. *Analytical and Bioanalytical Chemistry*, 405(1), 7-13. doi 10.1007/s00216-012-6473-x.

Rowe, P. G. (1994). *Design thinking*. MIT Press.

Savery, J. R. (2006). Overview of problem-based learning: Definitions and distinction. *Interdisciplinary Journal of Problem-based Learning*, /(1), 9-20.

Sendag, S., & Ferhan Odabasi, H. (2009). Effects of an online problem based learning course on content knowledge acquisition and critical thinking skills. Computers & Education, 53(1), 32-141; doi : 10.101 6/J.compedu. 2009.01.008.

Silverstein, L. B., & Layne, S. (2010). What is arts integration? Washington, DC: The Kennedy Center for the Performing Arts.

Solomon, B. (2013). *How Intermediate Grade Level Teachers' Culturally Relevant Beliefs and Practices Contribute to Mastery Achievement Levels for African American Students*. Dowling College.

Spurring African-American STEM Degree Completion. U.S. Department of Education. MARCH 16, 2016. https://www.ed.gov/news/press-releases/fact-sheet-spurring-african-american-stem-degree-completion

Stanford University Institute of Design. (2010). Bootcamp Bootleg. Bootcamp Bootleg, 47. http://dschool.stanford.edu/wp-content/uploads/2011/03/BootcampBootleg2010v2SLIM.pdf

Standley, J. (2008). Does music instruction help children learn to read? Evidence of a meta-analysis. Research in Music Education, 27, 17-32.

Sungur, S., & Tekkaya, C. (2006). Effects of problem-based learning and traditional instruction on self-regulated learning. *The Journal of Educational Research*, *99*(5), 307–320.

Tai R, Liu C, Maltese A, Fan X (2006) Planning early for careers in science. Science 312:1143-1144. doi :10.1126/science. 11 28690

Tan, E., & Calabrese Barton, A. (2018). Towards critical justice: Exploring intersectionality in community-based STEM-rich making with youth from non-dominant communities. *Equity & Excellence in Education: University of Massachusetts School of Education Journal*, *51*(1), 48–61.

Tapia, M. (1996). *The Attitudes toward Mathematics Instrument.* Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Tuscaloosa, AL, November 6-8, 1996).

Thomas, J. W. (2000). A review of research on Project-Based Learning (pp. 1-48). Retrieved from http://www.bie.org/index.php/site/RE/pbl_research/29.

Tomljenović, Z., & Tatalović Vorkapić, S. (2020). Constructivism in Visual Arts Classes. *Center for Educational Policy Studies Journal*, *10*(4), 13–32.

Turner S, Ireson G (2010) Fifteen pupils' positive approach to primary school science: When does it decline? Educ Stud 36: 1 19-141. doi: 10. 1080/0305569090314

U.S. Department of Education. (2016, March 16). *Spurring African-American STEM Degree Completion.* https://www.ed.gov/news/press-releases/fact-sheet-spurring-african-american-stem-degree-competion

U.S. Department of Education. (2016, September 22). *Supporting Science, Technology, Engineering and Mathematics (STEM) Success among African American Students.* White House Initiative on Educational Excellence for African Americans. https://sites.ed.gov/whblackeducation/files/2016/09/ STEM-Fact-Sheet-9.22.16.pdf

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Whisman, A. & Hixson, N. (2012). A cohort study of arts participation and academic\ performance. Charleston, WV: West Virginia Department of Education.

Wyse, A. E., Stickney, E. M., Butz, D., Beckler, A., & Close, C. N. (2020). *The Potential Impact of COVID‐19 on Student Learning and How Schools Can Respond*. Educational Measurement, Issues and Practice, 10.1111/emip.12357. https://doi.org/10.1111/emip.12357

Appendices

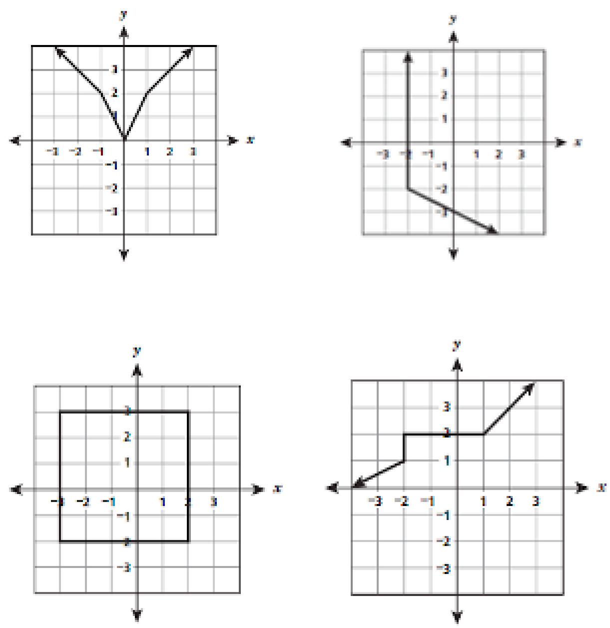
Appendix A - Functions Pre/Post Test

Name____ Hr________

1. Find the unit rate.

45 books in 5 shelves = ___________ books per shelf.

2. Which graph represents a function?



3. The table below shows how the amount Mary earns from yard work depends on the number of hours she works.

Time Spent (hours), x	3	4
Amount earned, y	$24	$32

How much money does Mary earn per hour? ________________

4. Darren and his friend save money for a vacation. Darren starts with \$120 in savings. Each week, Darren adds \$40 to his savings

Darren's friend also begins with \$120 in savings but saves at a faster rate than Darren.

Select all the equations that could represent the amount of money, y, in dollars, that Darren's friend saves for vacation in x weeks.*

$y = 120 + 50x$

$y = 120 + 35x$

$y = 120 + 39x$

$y = 75x + 120$

$y = 87x + 120$

$y = 21x + 120$

5. The four tables below show relationships in which the *x* values represent inputs and the *y* values represent the corresponding outputs.

Q

x	y
–2	–3
1	3
3	–3
5	3

R

x	y
–1	–5
2	4
3	7
4	10

S

x	y
–2	3
1	3
3	3
5	3

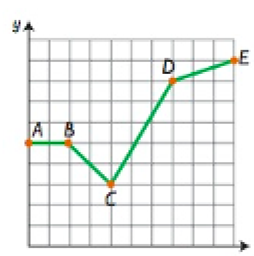
T

x	y
3	4
4	5
3	–4
4	–5

Which table represents a relationship that is **not** a function?

6.

Which section of the graph is decreasing?

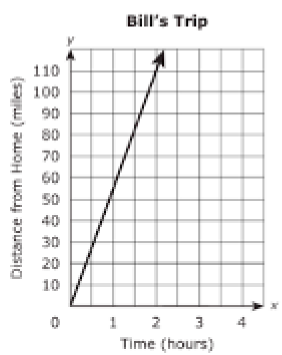


7. Which functions are not linear? Select all that apply.

a.   b. $y = 5 - x^2$ c. $-3x + 2y = 4$

d. $y = -5x - 2$ e. $y = 3x^2 + 1$ f. $y = x^3$

8. Bill drove his car at a constant speed while on a trip. Kevin drove his car at a different constant speed while on the same trip. The graph and the table show information about the trips Bill and Kevin took.



Kevin's Trip

Time (hours)	Distance from Home (miles)
0	0
2	90
3	135
5	225
6	270

Which sentence correctly compares the rates Bill and Kevin drove on their trip?

Bill drove at a rate that was 10 miles per hour slower than the rate Kevin drove.

Bill drove at a rate that was 10 miles per hour faster than the rate Kevin drove.

Bill drove at a rate that was 20 miles per hour slower than the rate Kevin drove.

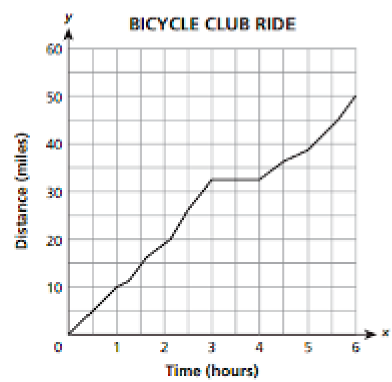
Bill drove at a rate that was 20 miles per hour faster than the rate Kevin drove.

9. Charlie is planning a party for 15 people. He finds a location that charges an initial fee of \$20 plus \$25 per person. What is the total cost of the party?

10. Consider the function: y = 2x + 5 What will be the value of the function when x = 10?

a. 52 b. 215 c. 25 d. 15

11. A bicycle club went on a six-hour ride. The graph below shows the relationship between the number of hours spent on the trails and the number of miles traveled.



Which statement best interprets information provided by the graph?

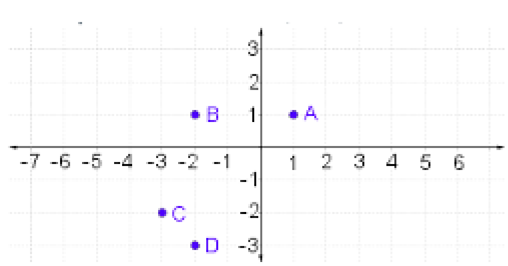
The club members rode at a constant speed for the entire ride.

The club members stopped for a rest during their ride.

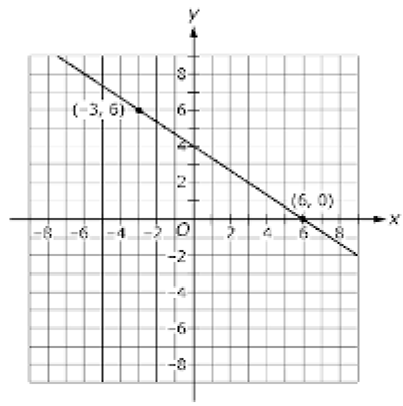
The number of miles traveled increased continuously throughout the ride.

The number of miles traveled increased some of the time and decreased some of the time.

12. Which point is located at (-3, -2)?



13. Which of the following equations shows the correct slope of the line whose graph is shown in the coordinate plane below?



$$y = -\frac{3}{2}x + 4 \qquad y = -\frac{2}{3}x + 4$$

$$y = \frac{2}{3}x + 6 \qquad y = \frac{3}{2}x + 6$$

14. Which of the input-output tables represents a function? Select **each** answer.

Input	Output
1	4
5	6
5	1
10	8

Input	Output
1	4
1	6
5	5
8	10

Input	Output
1	4
8	6
5	1
10	5

Input	Output
1	4
10	6
5	5
8	1

Input	Output
1	4
8	6
5	10
1	5

15.

The table shown below is generated using $f(x) = 3x + 7$.

x	f(x)
1	10
5	?
8	31
9	34

Find the missing value: $f(5) = ?$

Appendix B – Attitudes Toward Mathematics Inventory Survey

Self-Confidence Items

Item Number	**Item Statement**
9	**Mathematics is one of my most dreaded subjects.**
10	**My mind goes blank and I am unable to think clearly when working with mathematics.**
11	**Studying mathematics makes me feel nervous.**
12	**Mathematics makes me feel uncomfortable.**
13	**I am always under a terrible strain in a math class**.
14	**When I hear the word mathematics, I have a feeling of dislike.**
15	**It makes me nervous to even think about having to do a mathematics problem.**
16	Mathematics does not scare me at all.
17	I have a lot of self-confidence when it comes to mathematics.
18	I am able to solve mathematics problems without too much difficulty.
19	I expect to do fairly well in any math class I take.
20	**I am always confused in my mathematics classes.**
21	**I feel a sense of insecurity when attempting mathematics.**
22	I learn mathematics easily.
40	I believe I am good at solving math problems.

Bolded items are negatively stated

Value Items

Item Number	**Item Statement**
1	Mathematics is a very worthwhile and necessary subject
2	I want to develop my mathematical skills.
4	Mathematics helps develop the mind and teaches a person to think.
5	Mathematics is important in everyday life.
6	Mathematics is one of the most important subjects for people to study.
7	High school math courses would be very helpful no matter what I decide to study.
8	I can think of many ways that I use math outside of school.
35	I think studying advanced mathematics is useful.
36	I believe studying math helps me with problem solving in other areas.
39	A strong math background could help me in my professional life.

Enjoyment Items

Item Number	Item Statement
3	I get a great deal of satisfaction out of solving a mathematics problem.
24	I have usually enjoyed studying mathematics in school.
25	**Mathematics is dull and boring.**
26	I like to solve new problems in mathematics.
27	I would prefer to do an assignment in math than to write an essay.
29	I really like mathematics.
30	I am happier in a math class than in any other class.
31	Mathematics is a very interesting subject.
37	I am comfortable expressing my own ideas on how to look for solutions to a difficult problem in math.
38	I am comfortable answering questions in math class.

Bolded items are negatively stated

Motivation Items

Item Number	Item Statement
23	I am confident that I could learn advanced mathematics.
28	**I would like to avoid using mathematics in college.**
32	I am willing to take more than the required amount of mathematics.
33	I plan to take as much mathematics as I can during my education.
34	The challenge of math appeals to me.

Bolded items are negatively stated

Appendix C – Project-based Learning Activity

Description
The students will design and organize the needed components for creating a team for this year's St. Jude Mathathon.
Q2: Recruiting sponsors (advertising) [Could create some kind of recruiting and advertising materials]
Q3: calculating individual fund-raising goals (how many students, how many sponsors per student, how much per question?)
Q4: calculate a class goal to earn a special prize from the Mathathon (how much do we need to raise to hit a goal and how do we do that with our sponsors?) [Could create some kind of progress graphic to track actual fundraising for the class]

Key standards
8.F.A.1 - Understand that a function is a rule that assigns to each input exactly one output. The graph of a function is the set of ordered pairs consisting of an input and the corresponding output.

8.F.A.2 - Compare properties of two functions each represented in a different way (algebraically, graphically, numerically in tables, or by verbal descriptions).

8.F.A.3 - Interpret the equation $y = mx + b$ as defining a linear function, whose graph is a straight line; give examples of functions that are not linear.

8.F.B.4 - Construct a function to model a linear relationship between two quantities. Determine the rate of change and initial value from a description of the relationship or from two (x,y) values, including reading these from a table or graph. Interpret the rate of change and initial value of a linear function in terms of the situation it models, and in terms of its graph or a table of values.

8.B.F.5 - Describe qualitatively the function relationship between two quantities by analyzing a graph (e.g., where the function is increasing or decreasing, linear or nonlinear). Sketch a graph that exhibits the qualitative features of a function that has been described verbally.

Key Vocabulary
Function, initial value, input, output, linear function, nonlinear function, unit rate, domain, range

Literacy Skills

- Informational writing
- Engaging in collaborative conversations

Success Skills

- Collaboration
- Communication

Project Milestones

Milestone 1: What is a Function? "Build a machine"

Estimated Duration

3–5 days

Assessment(s)

What competencies should the students have when they finish this activity?
How do we measure student understanding of those competencies? (formative)

Key Student Question

What is the relationship between input and output?

Description

Capture students' need-to-know (before they begin the project) questions about the topic and the project.

Ask questions such as these to prompt student thinking about what they know and need to know:

Form project teams: Create teams of three or four students (Three is ideal). Mix students according to their leadership, technology, and communication skills. Assign roles to each student: Leader (keeps everyone active and contributing), Recorder (writes down whatever the group decides), Question Keeper (makes sure that the group keeps answering the right question. The only team member who may ask a question of the teacher)

Students identify a machine that they want to build where the user inputs something and expects an output. How does the machine create that output? Is it possible to get the wrong output from your machine? What might cause an incorrect output?

EI: The human vending machine. A chain of three students form the machine. There is a cooler with drinks in front of the middle student.

STEP 1: An outside user hands the first student a Monopoly dollar. The first tells the second “one drink.” The second pulls one drink out and hands it to the third student who delivers the drink to the outside user. Repeat by handing the first student 3 Monopoly dollars…how many drinks are delivered? 5 dollars…how many drinks are delivered?

STEP 2: Same process, except now there are two types of drinks; water and Gatorade. Water costs $0.50 and Gatorade $1. What happens when the machine is fed $1.50? $2? $5?

Record the results in a table. How many drinks are given out? How many dollars are collected? Where might a mistake be made in the process? How can that be corrected?

Expectation for Final Product
The final product for this portion of the project consists of…students must demonstrate their machine and how it works. They should input at least 3 amounts in STEP 1 and 3 amounts in STEP 2.

Milestone 2: How do we recruit Sponsors?

Estimated Duration
2 weeks

Assessment(s)
What competencies should the students have when they finish this activity?
How do we measure student understanding of those competencies? (formative)

Key Student Question
What do we need to recruit sponsors? How does sponsorship work? How much can we raise?

Description
Capture students’ need-to-know questions about the topic and the project.
Ask questions such as these to prompt student thinking about what they know and need to know:
What is a mathathon? How does the mathathon work? What is the purpose? Who does it raise money for? What do we do to raise money for this?

Form project teams: Create teams of three or four students. Mix students according to their leadership, technology, and communication skills.

Getting sponsors
Goal - Create a pledge form to ask people to pledge money.
Who do we ask to sponsor us? How does sponsor pledges work?
IE:

- Mom pledges - 10 cents per problem answered correctly 0.10x=y
- Grandma pledges - 5 cents per problem answered correctly but will give you $5 to start 0.05x + 5=y
- Aunt pledges - 15 cents per problem attempted 0.15x=y

Projecting how much money you can raise: if you answer some of the questions, if you answer all of the questions. What happens if you can convince your sponsors to give you more for each problem? Graph the progress of each of your pledges. Which pledge earns you the most the fastest? (comparing slopes)

Expectation for Final Product
The final product for this portion of the project consists of…
Create a pledge form to ask people to pledge money. Create some type of material (video, infographic, brochure) to use to explain to your possible sponsors how much you can raise and what the money goes to. Go and begin getting sponsors!

Milestone 3: How much money can we raise as a team of three?

Estimated Duration
2 weeks

Assessment(s)
Need-to-know questions (whole group)

Key Student Question
What do we?

Description
Capture students need-to-know questions about the topic and the project.
Ask questions such as these to prompt student thinking about what they know and need to know:

How many pledges have you gotten? Can you graph how much money you would raise from each sponsor if you answered 10 problems? 20 problems? All of the problems?

Form project teams: In your teams, pass your graphs to a teammate. That teammate uses the graphs to figure out what the function is for each sponsor.

Now, we know how much each member of the team can earn TOTAL from all of their sponsors combined. Compare the earnings of each team member by looking at the earning slopes together on one graph.

Expectation for Final Product
The final product for this portion of the project consists of…present to the class as a team, how much they can raise. The teams compete to see who can raise the most money. Teams have to graphically represent their potential earnings. Each team will present their findings (pledges, graphs, videos, etc.) to the principal and she will grade the students based on accuracy and creativity. A donation of her choice will be given to the winning team.

Milestone 4: How do we raise enough money to get the special prize as a class?

Estimated Duration
2 weeks

Assessment(s)
Need-to-know questions (whole group)

Key Student Question
How much do we have to raise as a class? As a school to get the special prize from St. Jude?

Description
Capture students' need-to-know questions about the topic and the project.

Ask questions such as these to prompt student thinking about what they know and need to know:
How much money raising potential does my team have? How did we combine the earnings of each team member to determine that potential?

Form project teams: Create teams of three or four students. Mix students according to their leadership, technology, and communication skills.

Expectation for Final Product

The final product for this portion of the project consists of… Create and maintain a graphic tracker to keep track of how much the class is earning as they go through solving problems for the mathathon. Create a competition between the 8th grade math classes for who raises the most. The class that raises the most gets…? Each class creates a wrap-up video explaining how much they raised, what they learned about how a Mathathon works, and why it was important to do this. Throw a party for the whole 8th grade to celebrate the important work they just did.

Appendix D - Statistical Analysis for Renaissance STAR Exam

Pre-test T-test comparing Treatment and Control Groups

Control Group (0)

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.895073	**Pr < W**	0.0025
Kolmogorov-Smirnov	**D**	0.148639	**Pr > D**	0.0432
Cramer-von Mises	**W-Sq**	0.185081	**Pr > W-Sq**	0.0079
Anderson-Darling	**A-Sq**	1.137422	**Pr > A-Sq**	<0.0050

Treatment Group (1)

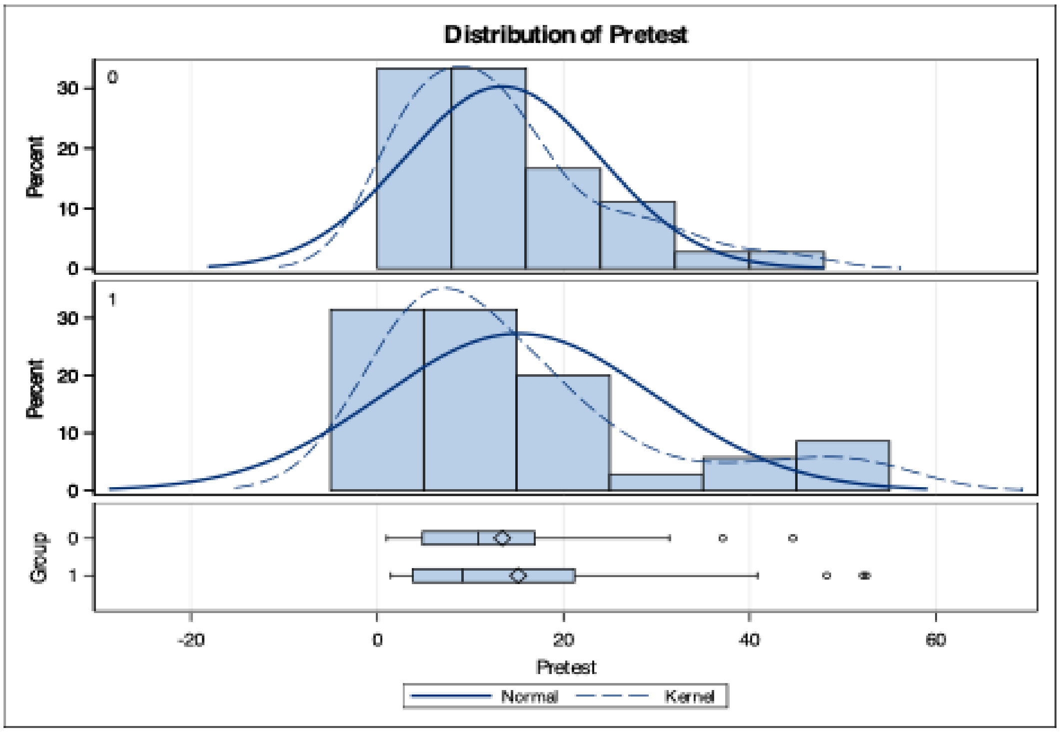
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.804525	**Pr < W**	<0.0001
Kolmogorov-Smirnov	**D**	0.177369	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	0.385543	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	2.363304	**Pr > A-Sq**	<0.0050

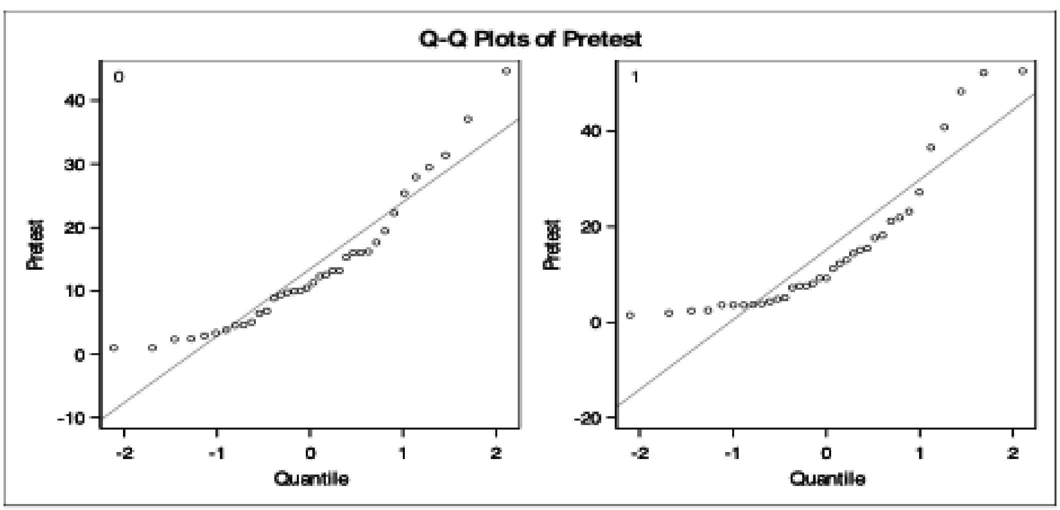
Group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		36	13.4394	10.5376	1.7563	1.0000	44.6364
1		35	15.1481	14.6514	2.4765	1.4545	52.5455
Diff (1-2)	**Pooled**		-1.7087	12.7319	3.0223		
Diff (1-2)	**Satterthwaite**		-1.7087		3.0361		

Group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		13.4394	9.8740	17.0048	10.5376	8.5469	13.7457
1		15.1481	10.1151	20.1810	14.6514	11.8511	19.1963
Diff (1-2)	**Pooled**	-1.7087	-7.7380	4.3207	12.7319	10.9166	15.2771
Diff (1-2)	**Satterthwaite**	-1.7087	-7.7784	4.3610			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	69	-0.57	0.5737
Satterthwaite	Unequal	61.65	-0.56	0.5756

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	34	35	1.93	0.0562

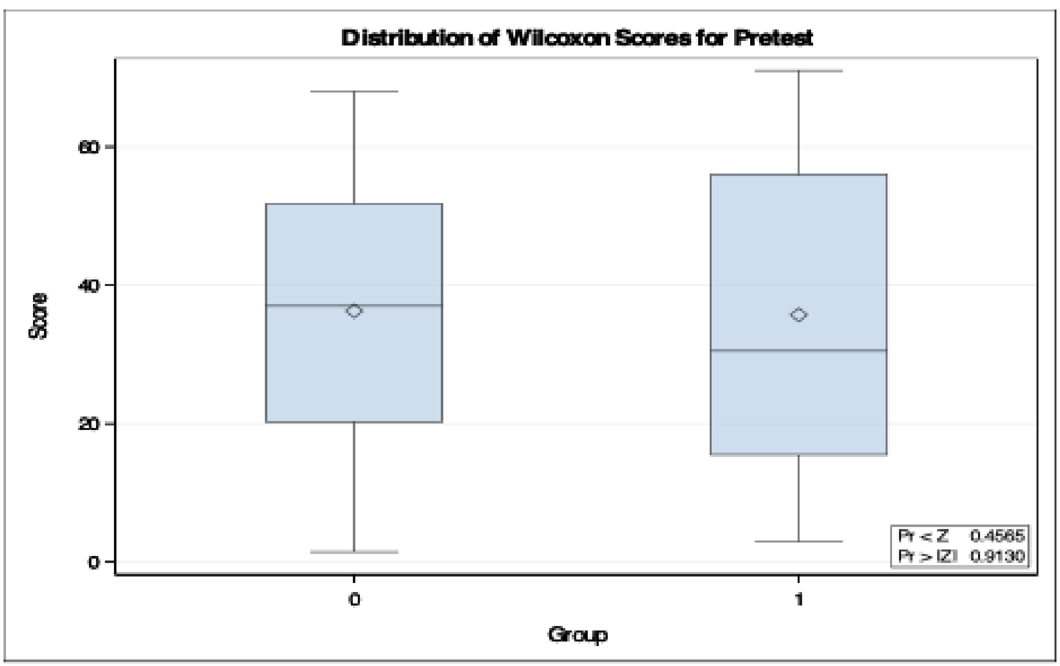




Wilcoxon Scores (Rank Sums) for Variable Pretest Classified by Variable Group					
Group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
0	36	1306.0	1296.0	86.937326	36.277778
1	35	1250.0	1260.0	86.937326	35.714286
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	Z	Pr < Z	Pr > |Z|	Pr < Z	Pr > |Z|
1250.000	-0.1093	0.4565	0.9130	0.4566	0.9133
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
0.0132	1	0.9084



ANCOVA results comparing Treatment and Control Groups on Teacher-generated assessment

Class Level Information		
Class	**Levels**	**Values**
Group	2	0 1

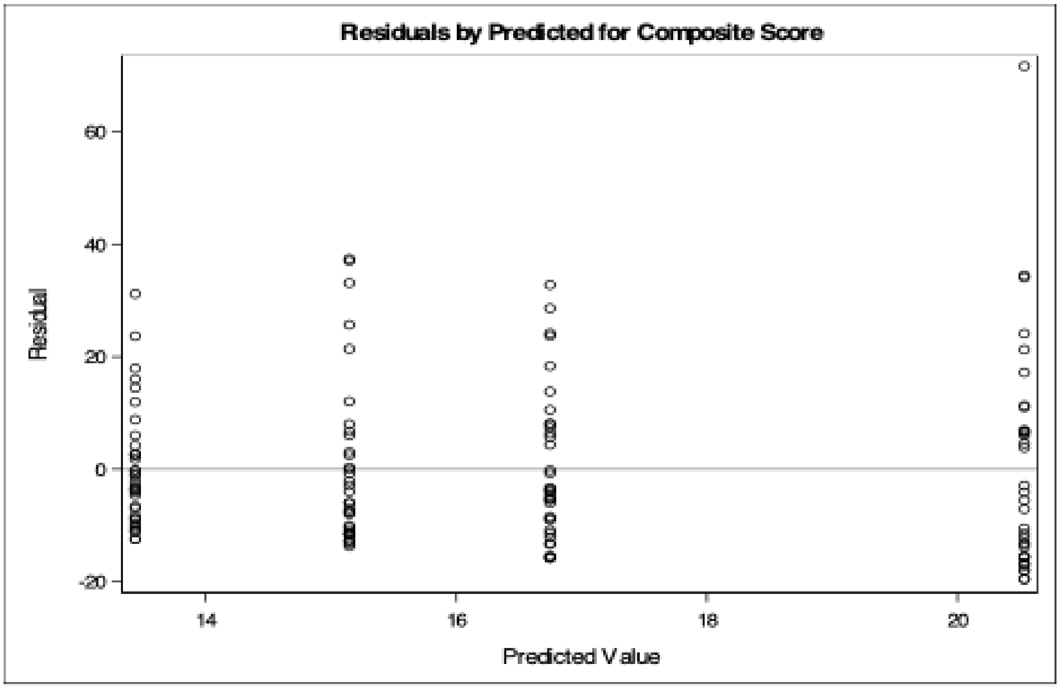
Number of Observations Read	1002
Number of Observations Used	142

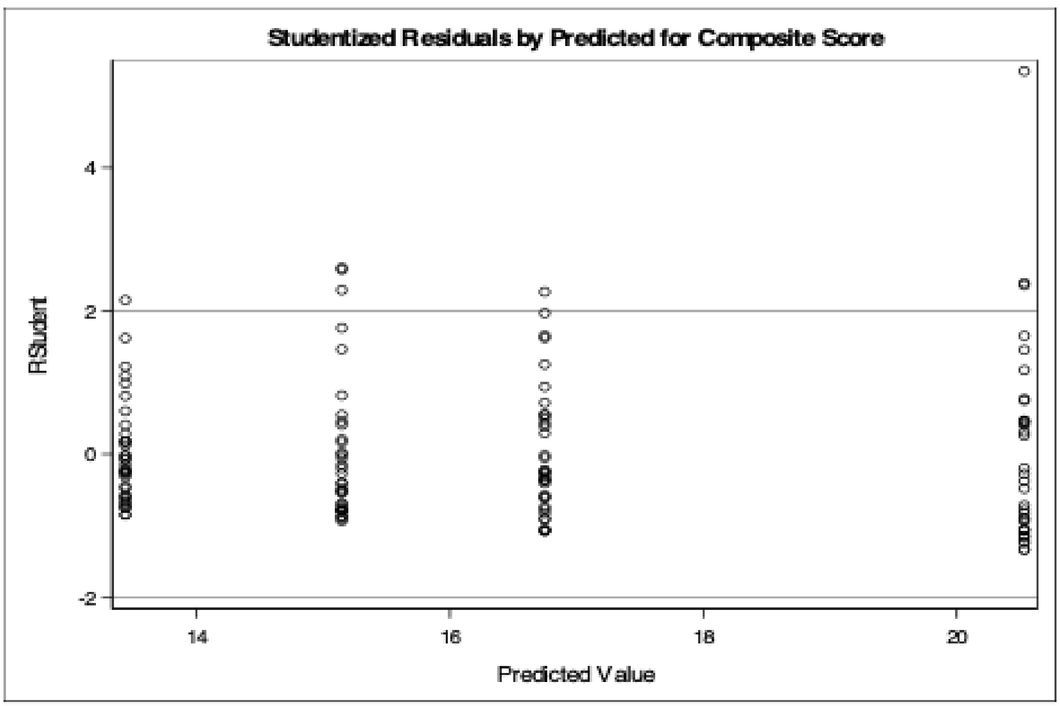
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	972.13610	324.04537	1.46	0.2283
Error	138	30634.57586	221.98968		
Corrected Total	141	31606.71197			

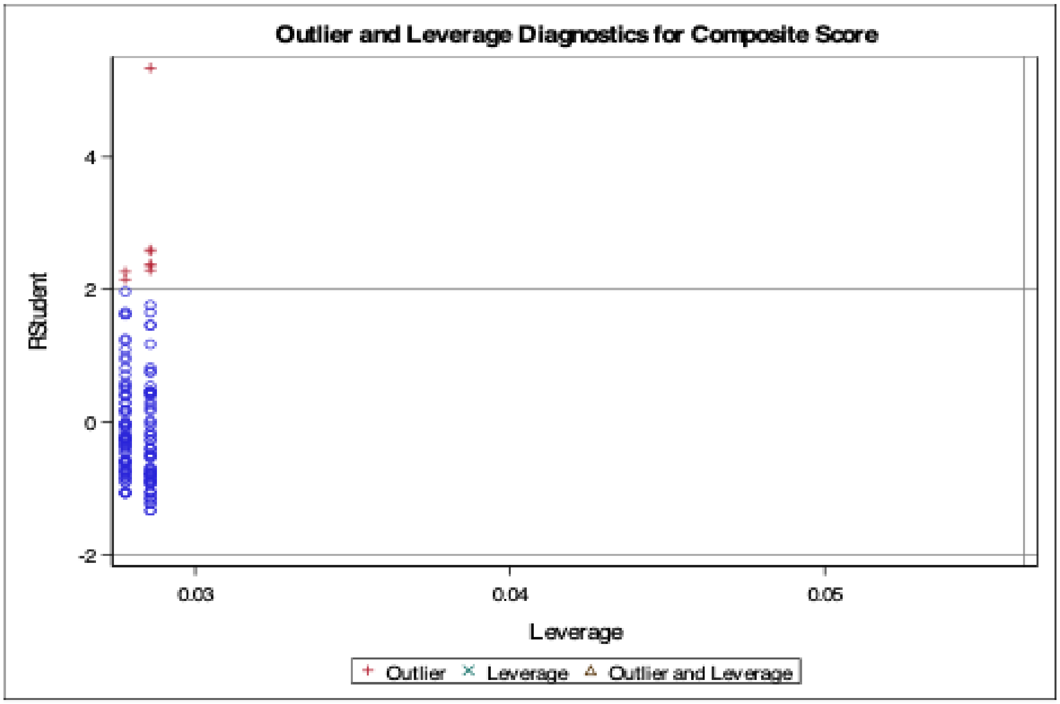
R-Square	Coeff Var	Root MSE	Composite Score Mean
0.030757	90.58711	14.89932	16.44750

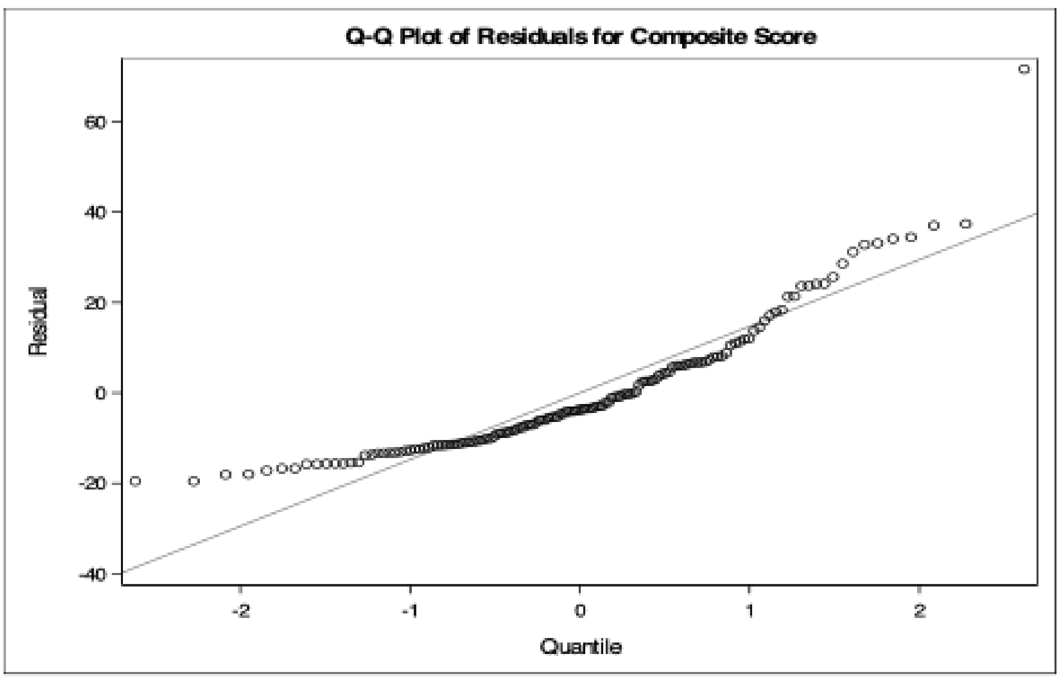
Source	DF	Type I SS	Mean Square	F Value	Pr > F
Group	1	267.7961323	267.7961323	1.21	0.2740
Pre-post	1	666.0859621	666.0859621	3.00	0.0855
Pre-post*Group	1	38.2540091	38.2540091	0.17	0.6787

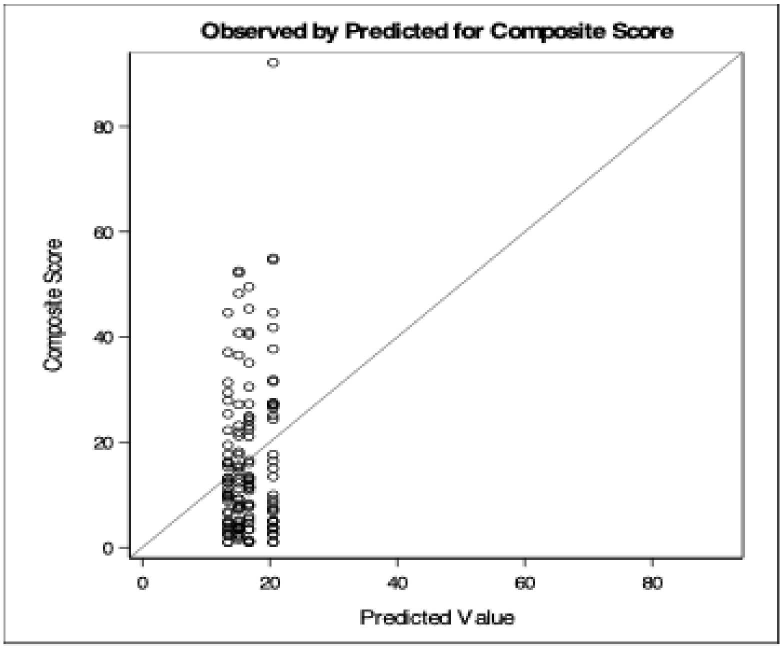
Source	DF	Type III SS	Mean Square	F Value	Pr > F
Group	1	51.8110614	51.8110614	0.23	0.6298
Pre-post	1	670.4574778	670.4574778	3.02	0.0845
Pre-post*Group	1	38.2540091	38.2540091	0.17	0.6787


Residuals by Predicted for Composite Score
Residual
60
40
20
0
-20
14
16
18
20
Predicted Value

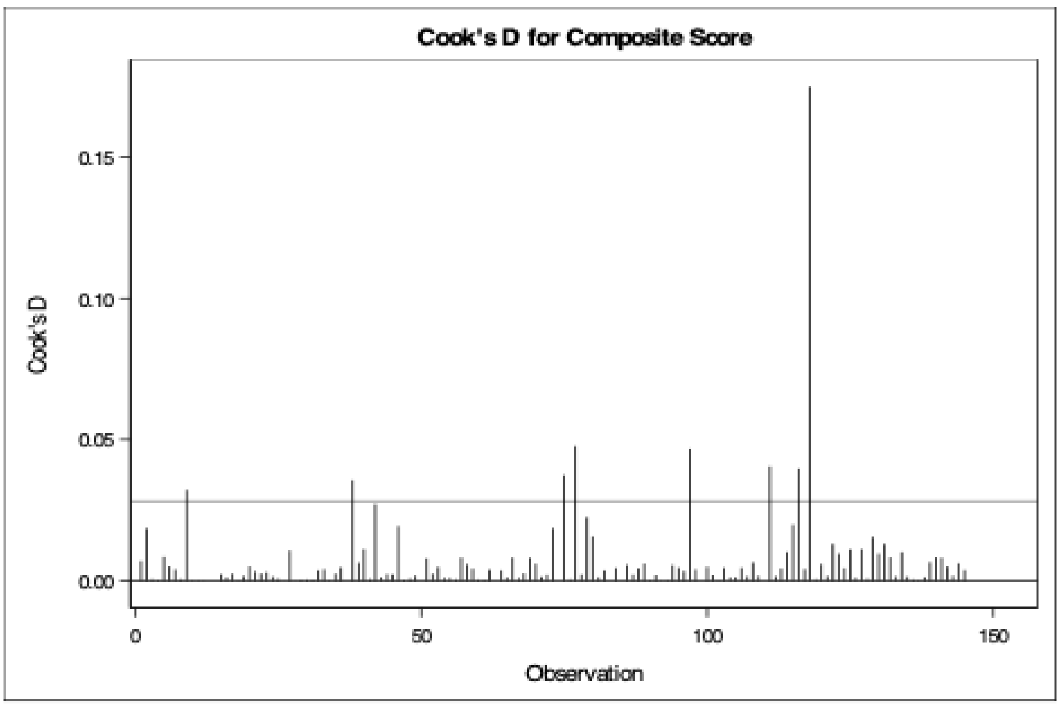

Studentized Residuals by Predicted for Composite Score
RStudent
4
2
0
-2
14
16
18
20
Predicted Value

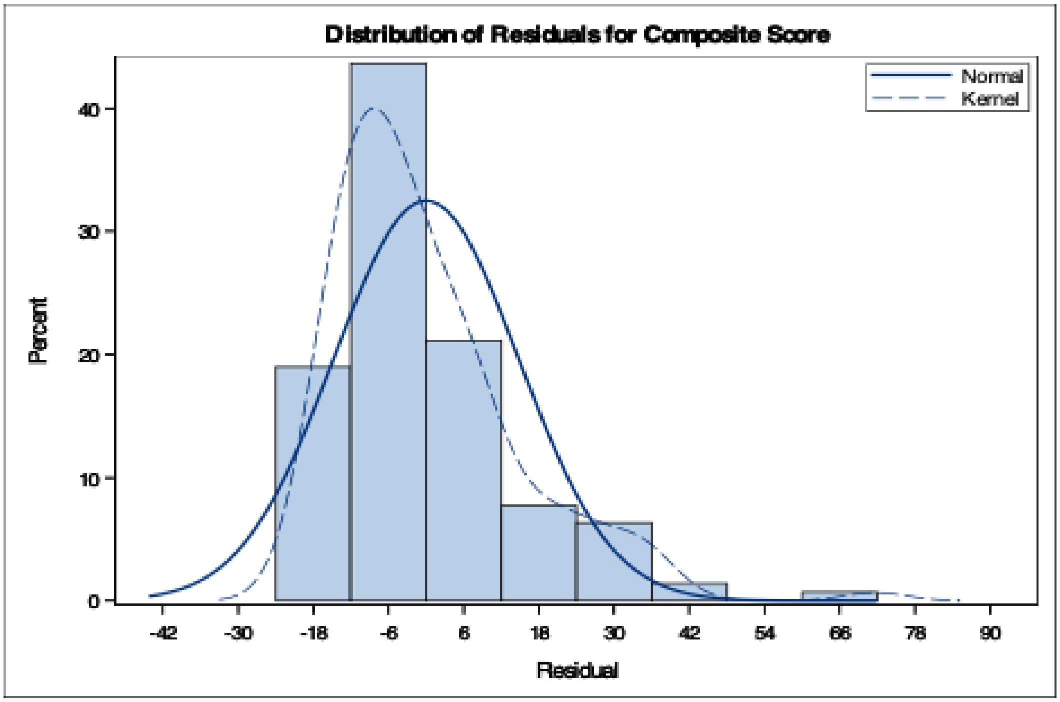

Outlier and Leverage Diagnostics for Composite Score
RStudent
4
2
0
-2
0.03
0.04
0.05
Leverage
Outlier
Leverage
Outlier and Leverage

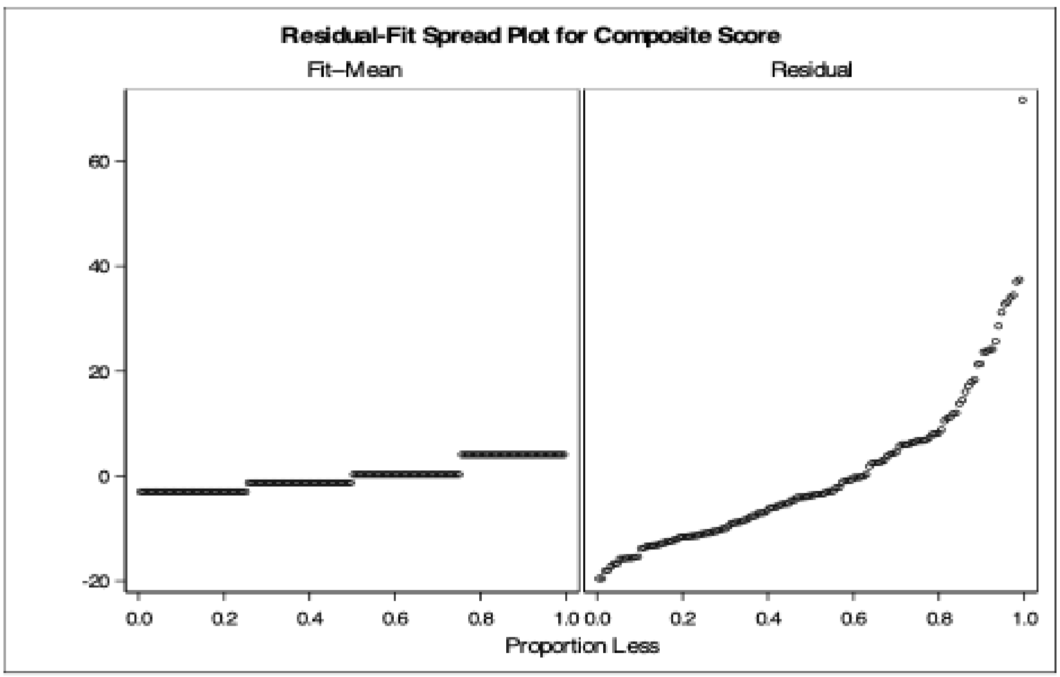

Q-Q Plot of Residuals for Composite Score
Residual
60
40
20
0
-20
-40
-2
-1
0
1
2
Quantile

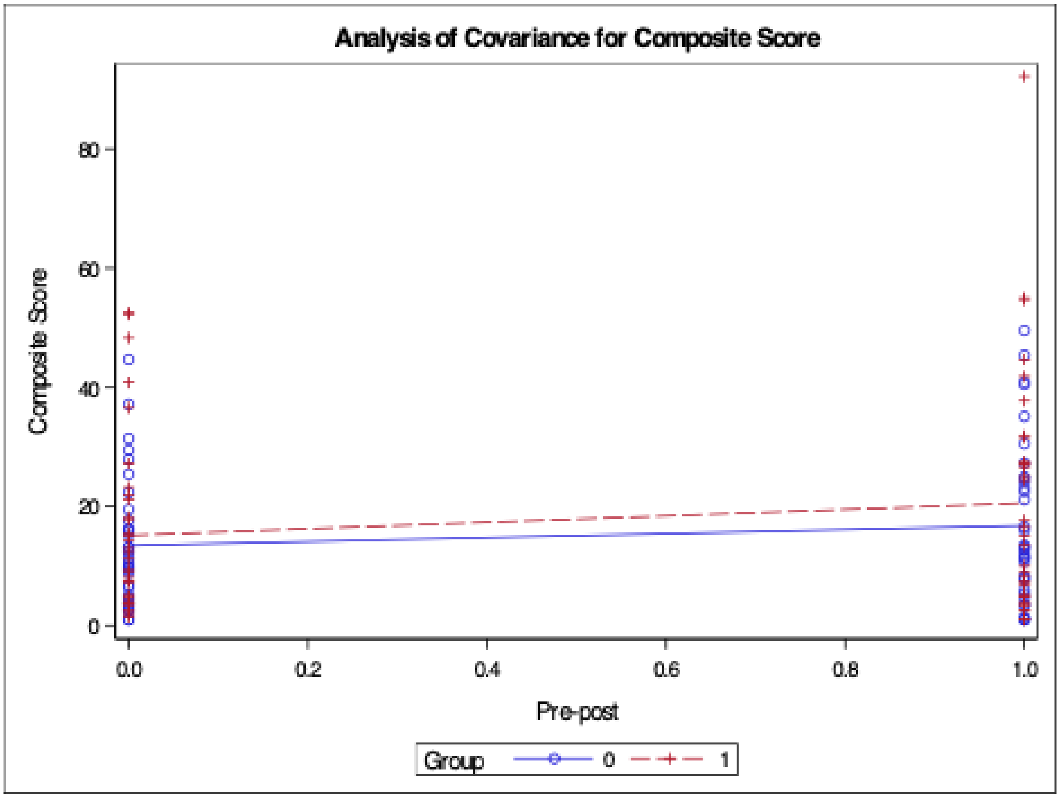

Observed by Predicted for Composite Score
Composite Score
0
20
40
60
80
Predicted Value

.

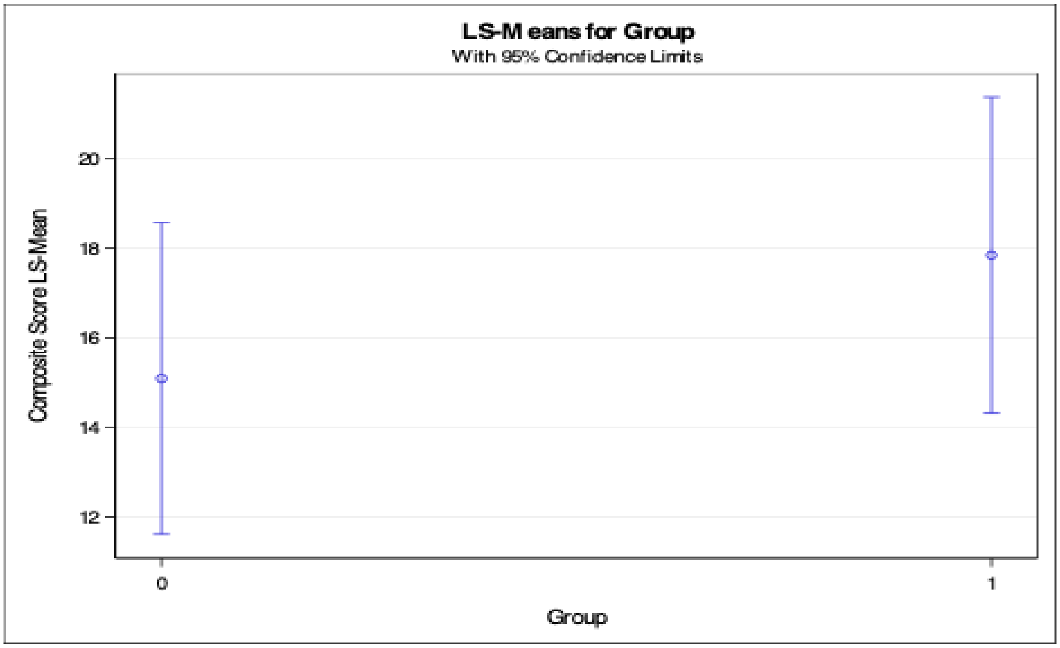

Cook's D for Composite Score
Cook's D
0.00
0.05
0.10
0.15
0
50
100
150
Observation


Distribution of Residuals for Composite Score
Normal
Kernel
Percent
0
10
20
30
40
-42
-30
-18
-6
6
18
30
42
54
66
78
90
Residual


Residual-Fit Spread Plot for Composite Score
Fit–Mean
Residual
-20
0
20
40
60
0.0
0.2
0.4
0.6
0.8
1.0
Proportion Less



Group	Composite Score LSMEAN	H0:LSMean1=LSMean2
		Pr > |t|
0	15.0934343	0.2740
1	17.8402597	


LS-Means for Group
With 95% Confidence Limits
Composite Score LS-Mean
20
18
16
14
12
0
1
Group

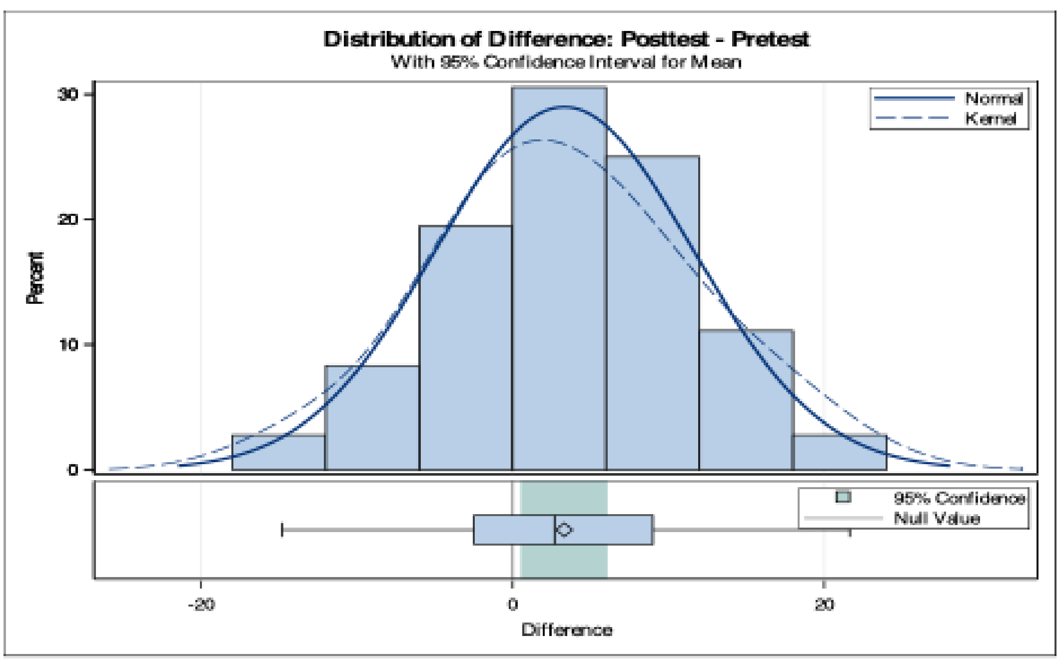
Pre-post Paired T-test Control Group on Renaissance STAR Exam

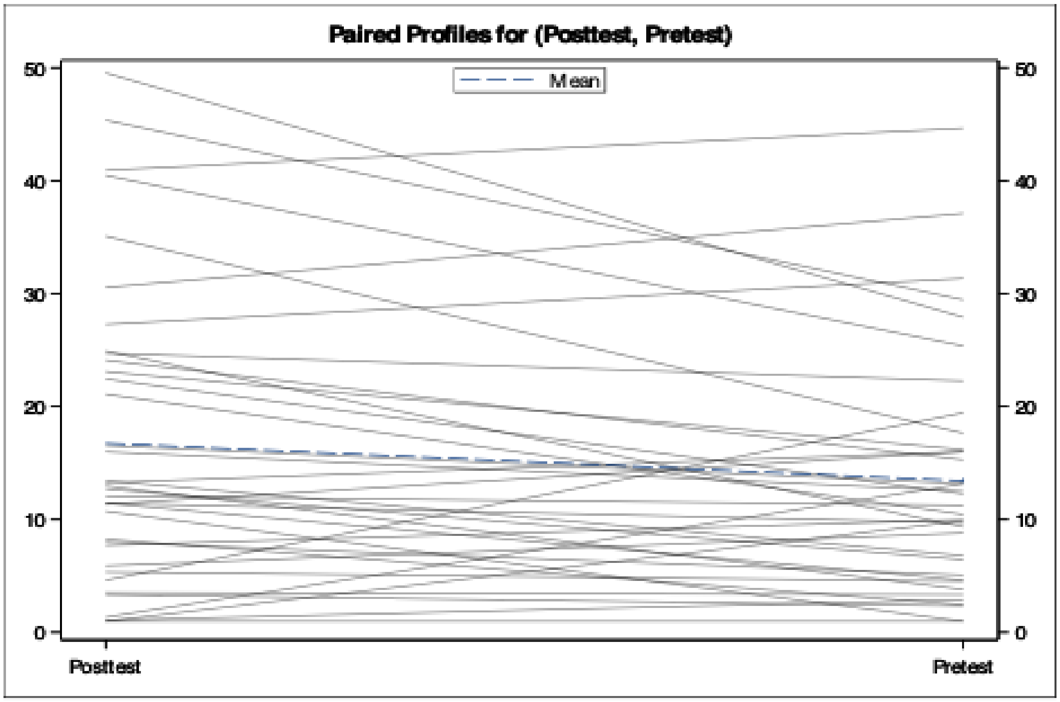
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.98994	**Pr < W**	0.9820
Kolmogorov-Smirnov	**D**	0.076259	**Pr > D**	>0.1500
Cramer-von Mises	**W-Sq**	0.025901	**Pr > W-Sq**	>0.2500
Anderson-Darling	**A-Sq**	0.171647	**Pr > A-Sq**	>0.2500

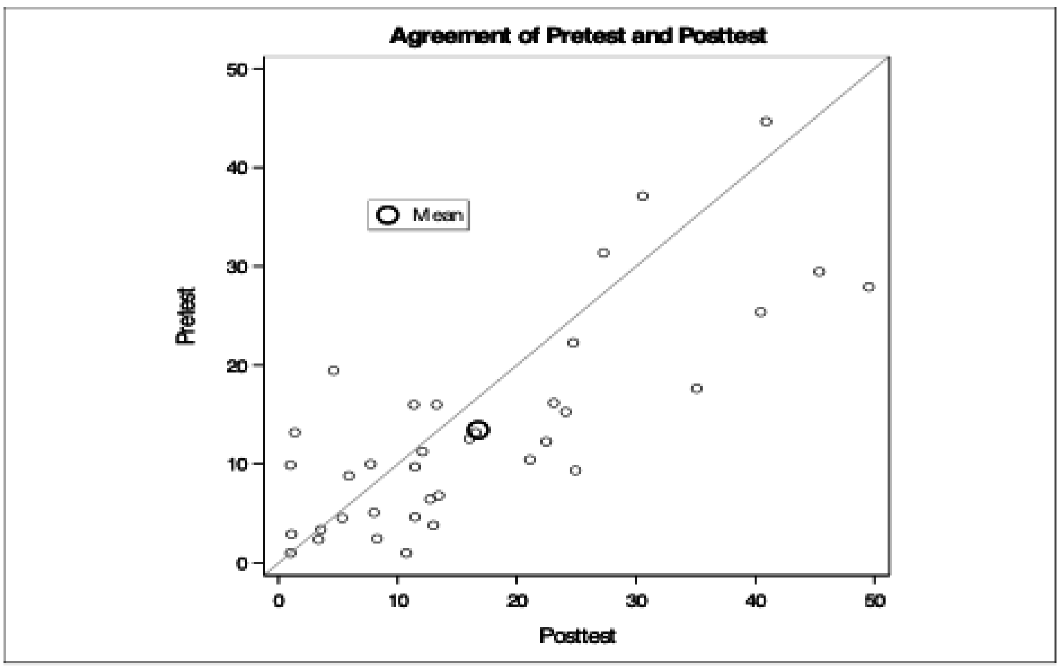
N	Mean	Std Dev	Std Err	Minimum	Maximum
36	3.3081	8.2589	1.3765	-14.8182	21.6364

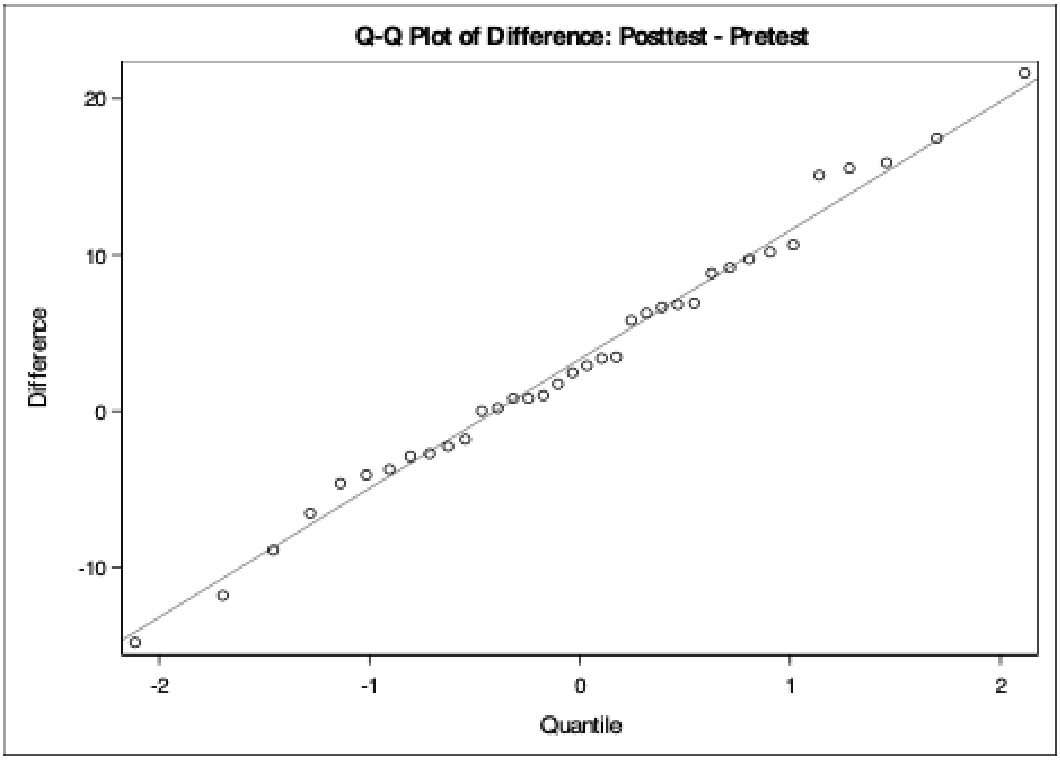
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
3.3081	0.5137	6.1025	8.2589	6.6986	10.7732

DF	t Value	Pr > |t|
35	2.40	0.0217




Paired Profiles for (Posttest, Pretest)
Mean
50
40
30
20
10
0
Posttest
Pretest


Agreement of Pretest and Posttest
Mean
Pretest
Posttest
0
10
20
30
40
50


Q-Q Plot of Difference: Posttest - Pretest
20
10
0
-10
Difference
-2
-1
0
1
2
Quantile

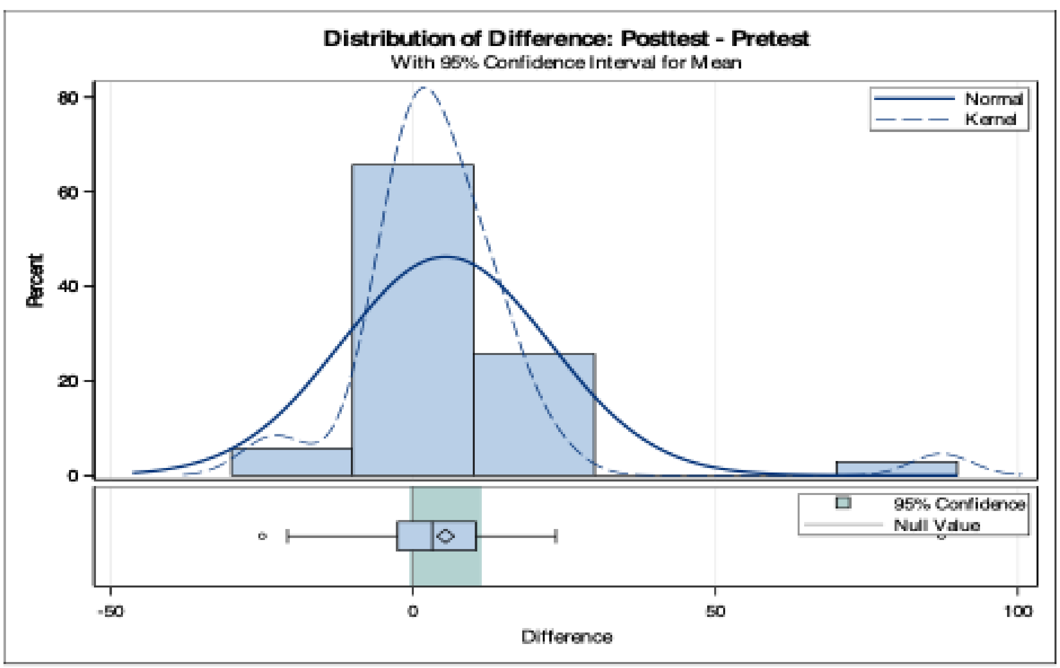
Pre-post Paired T-test Treatment Group on Renaissance STAR Exam

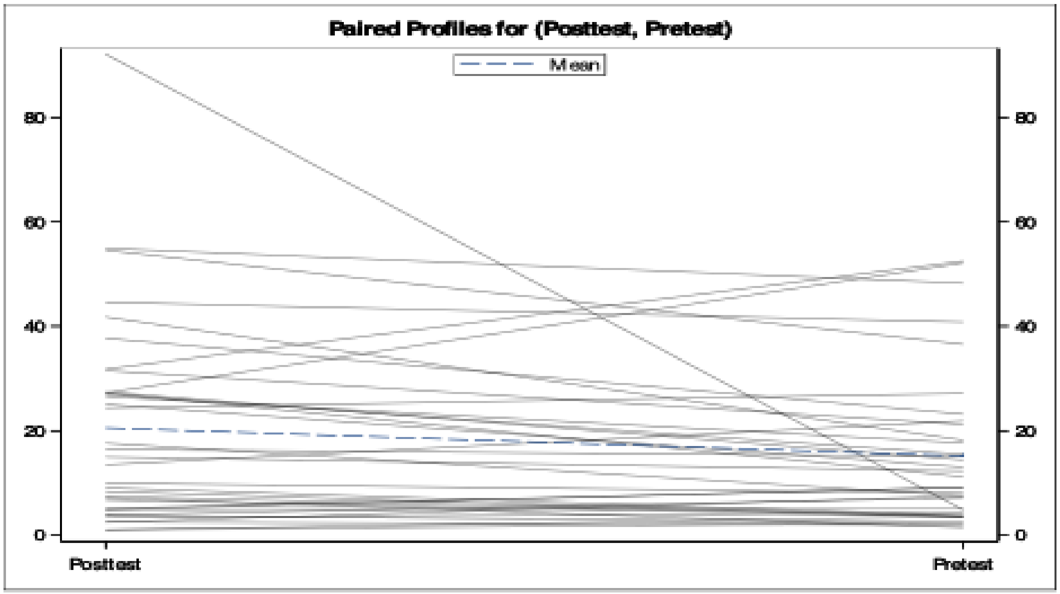
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.697559	**Pr < W**	<0.0001
Kolmogorov-Smirnov	**D**	0.186033	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	0.437361	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	2.633647	**Pr > A-Sq**	<0.0050

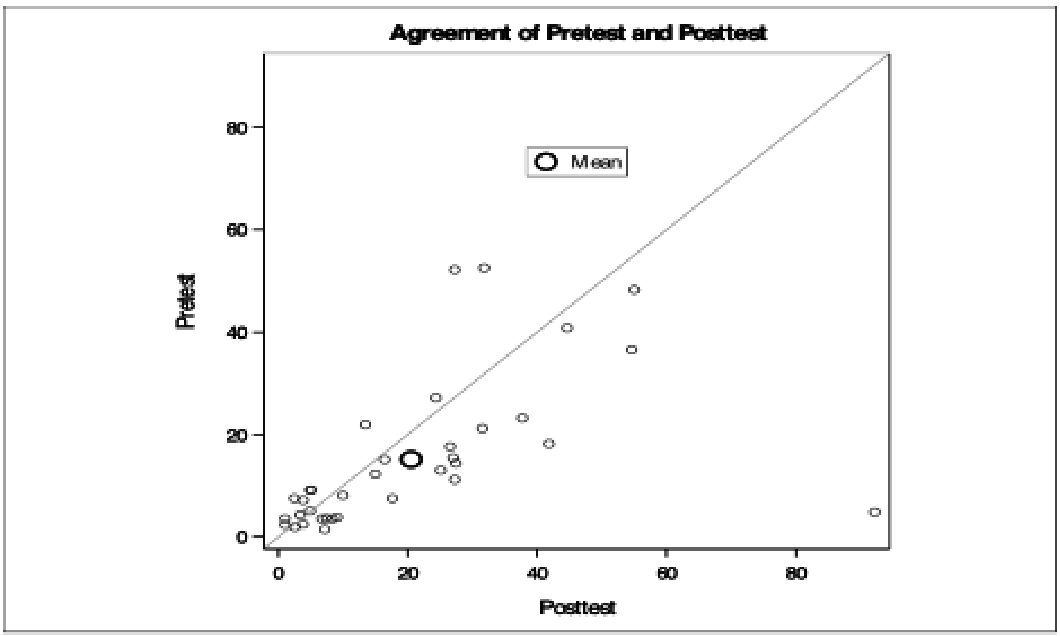
N	Mean	Std Dev	Std Err	Minimum	Maximum
35	5.3844	17.2423	2.9145	-24.9091	87.3636

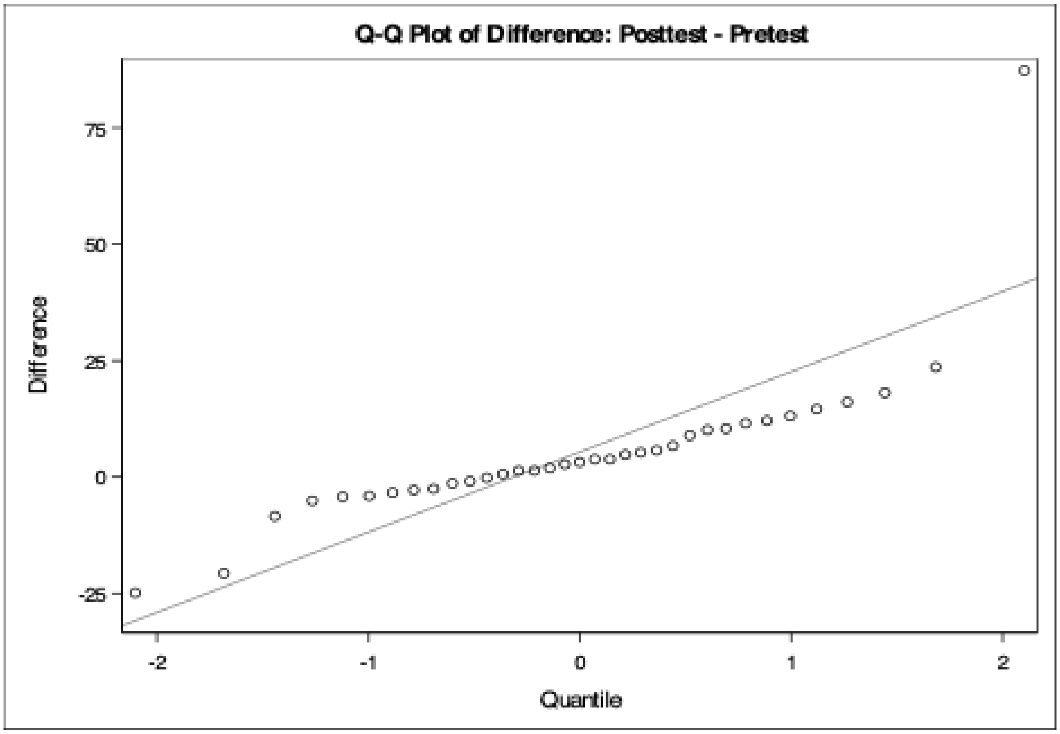
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
5.3844	-0.5385	11.3073	17.2423	13.9468	22.5909

DF	t Value	Pr > |t|
34	1.85	0.0734




Paired Profiles for (Posttest, Pretest)
Mean
0
20
40
60
80
Posttest
Pretest


Agreement of Pretest and Posttest
Mean
Pretest
Posttest
0
20
40
60
80



Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	1.847471	**Pr > |t|**	0.0734
Sign	**M**	5.5	**Pr >= |M|**	0.0895
Signed Rank	**S**	139.5	**Pr >= |S|**	0.0200

Post T-test comparing Treatment and Control Groups

Control Group (0)

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.907154	**Pr < W**	0.0054
Kolmogorov-Smirnov	**D**	0.181779	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	0.173742	**Pr > W-Sq**	0.0109
Anderson-Darling	**A-Sq**	1.063876	**Pr > A-Sq**	0.0079

Treatment Group (1)

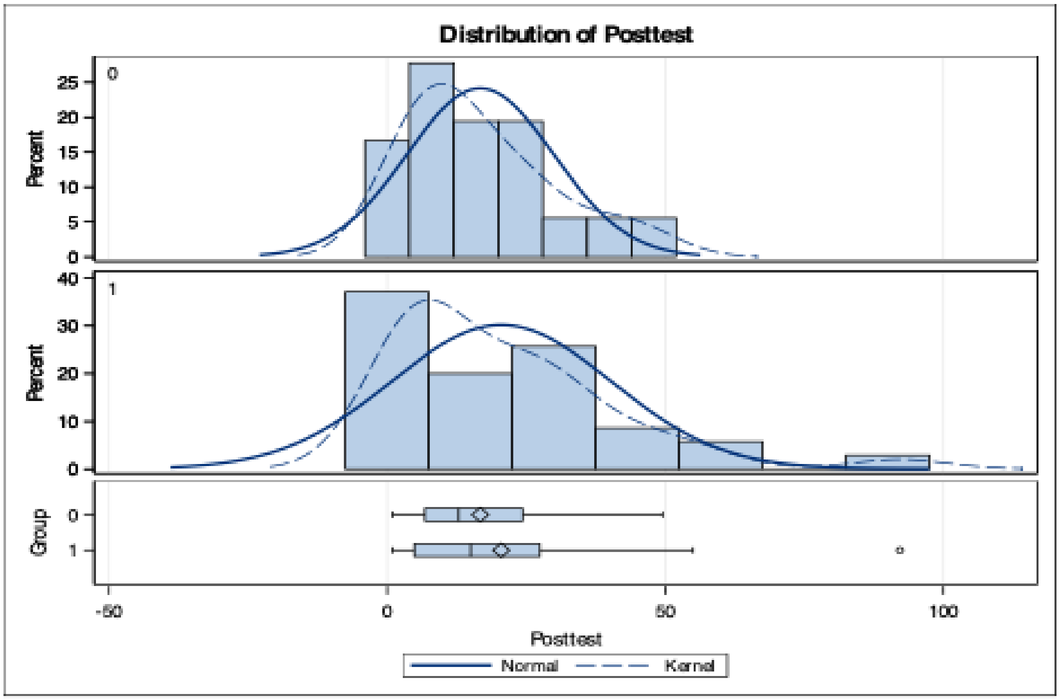
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.835372	**Pr < W**	0.0001
Kolmogorov-Smirnov	**D**	0.162083	**Pr > D**	0.0202
Cramer-von Mises	**W-Sq**	0.237934	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	1.528107	**Pr > A-Sq**	<0.0050

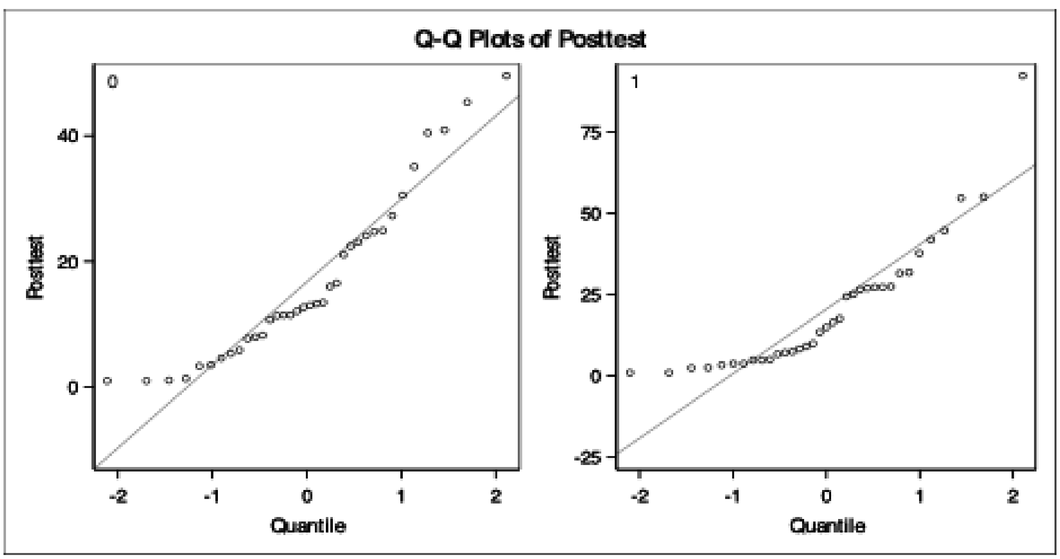
Group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		36	16.7475	13.2074	2.2012	1.0000	49.5455
1		35	20.5325	19.8111	3.3487	1.0000	92.1818
Diff (1-2)	**Pooled**		-3.7850	16.7892	3.9854		
Diff (1-2)	**Satterthwaite**		-3.7850		4.0074		

Group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		16.7475	12.2787	21.2162	13.2074	10.7123	17.2282
1		20.5325	13.7271	27.3378	19.8111	16.0247	25.9566
Diff (1-2)	**Pooled**	-3.7850	-11.7357	4.1657	16.7892	14.3954	20.1455
Diff (1-2)	**Satterthwaite**	-3.7850	-11.8037	4.2337			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	69	-0.95	0.3456
Satterthwaite	Unequal	59.025	-0.94	0.3488

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	34	35	2.25	0.0193

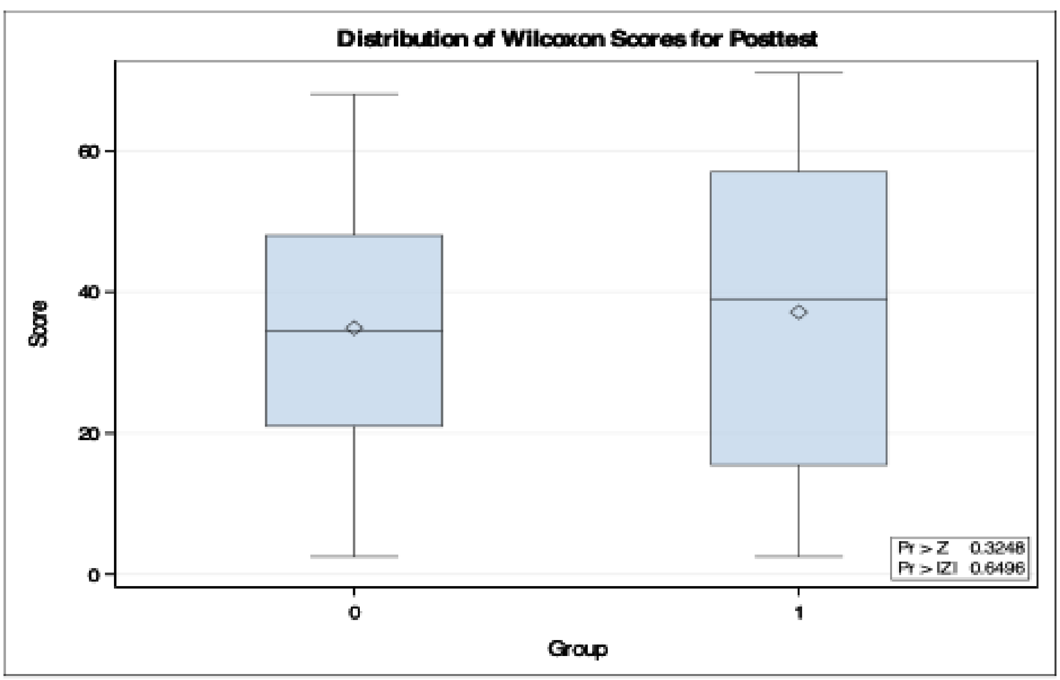




Wilcoxon Scores (Rank Sums) for Variable Posttest Classified by Variable Group					
Group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
0	36	1256.0	1296.0	86.934409	34.888889
1	35	1300.0	1260.0	86.934409	37.142857
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	Z	Pr > Z	Pr > |Z|	Pr > Z	Pr > |Z|
1300.000	0.4544	0.3248	0.6496	0.3255	0.6510
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
0.2117	1	0.6454



Appendix E - Statistical Analysis for Teacher-generated Assessment

Pre-test T-test comparing Treatment and Control Groups

Control Group (0)

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.912474	**Pr < W**	0.0076
Kolmogorov-Smirnov	**D**	0.200022	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	0.264111	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	1.444921	**Pr > A-Sq**	<0.0050

Treatment Group (1)

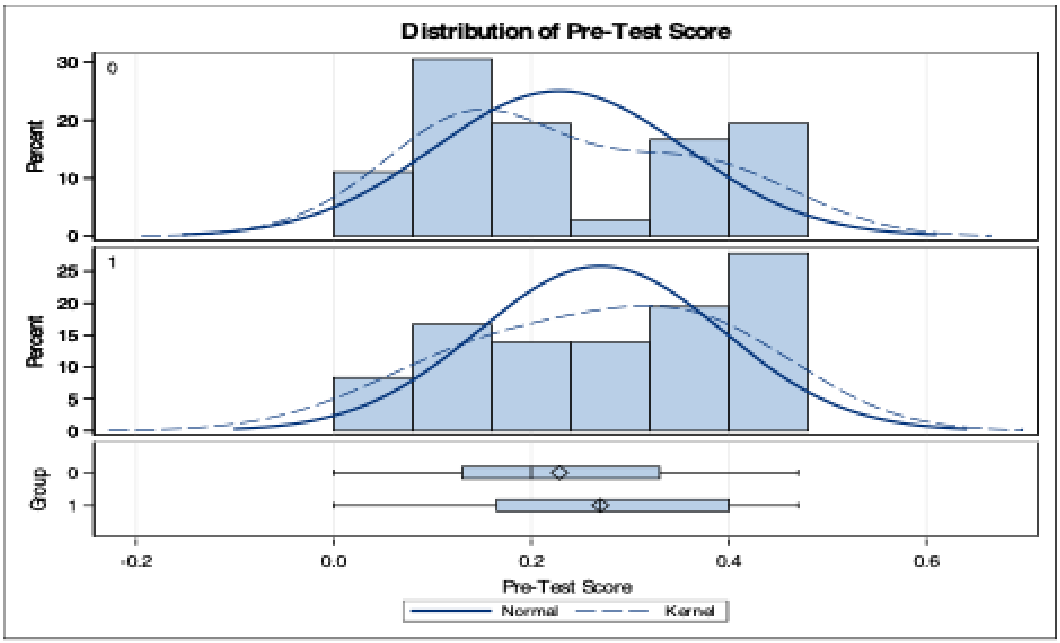
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.940928	**Pr < W**	0.0543
Kolmogorov-Smirnov	**D**	0.158425	**Pr > D**	0.0219
Cramer-von Mises	**W-Sq**	0.139714	**Pr > W-Sq**	0.0322
Anderson-Darling	**A-Sq**	0.857572	**Pr > A-Sq**	0.0246

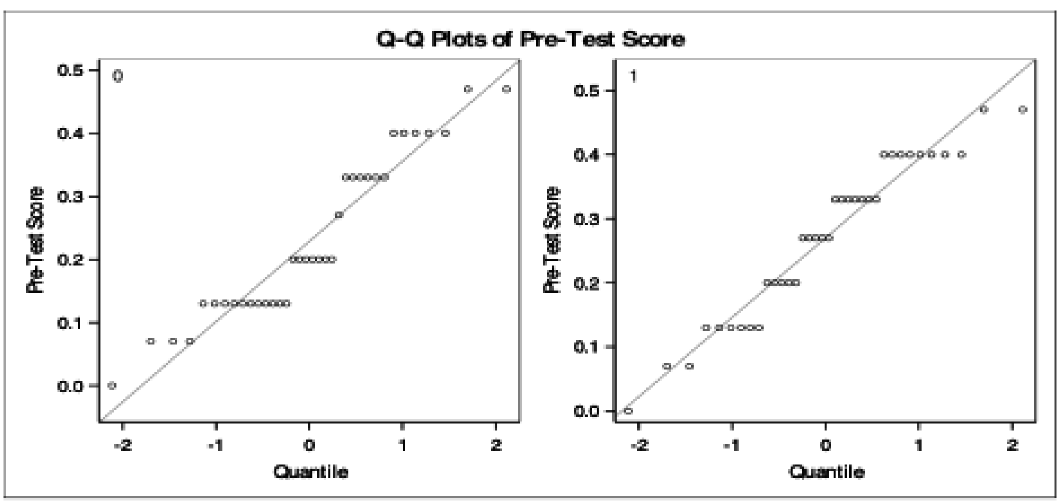
Group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		36	0.2286	0.1273	0.0212	0	0.4700
1		36	0.2700	0.1237	0.0206	0	0.4700
Diff (1-2)	**Pooled**		-0.0414	0.1255	0.0296		
Diff (1-2)	**Satterthwaite**		-0.0414		0.0296		

Group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		0.2286	0.1855	0.2717	0.1273	0.1033	0.1661
1		0.2700	0.2282	0.3118	0.1237	0.1003	0.1613
Diff (1-2)	**Pooled**	-0.0414	-0.1004	0.0176	0.1255	0.1077	0.1504
Diff (1-2)	**Satterthwaite**	-0.0414	-0.1004	0.0176			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	70	-1.40	0.1662
Satterthwaite	Unequal	69.942	-1.40	0.1662

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	35	35	1.06	0.8653





Wilcoxon Scores (Rank Sums) for Variable Pre-Test Score Classified by Variable Group					
Group	**N**	**Sum of Scores**	**Expected Under H0**	**Std Dev Under H0**	**Mean Score**
0	36	1190.50	1314.0	87.428039	33.069444
1	36	1437.50	1314.0	87.428039	39.930556
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	**Z**	**Pr < Z**	**Pr > |Z|**	**Pr < Z**	**Pr > |Z|**
1190.500	-1.4069	0.0797	0.1595	0.0819	0.1638
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	**DF**	**Pr > ChiSq**
1.9954	1	0.1578

ANCOVA results comparing Treatment and Control Groups on Teacher-generated assessment

Class Level Information		
Class	**Levels**	**Values**
Group	2	0 1

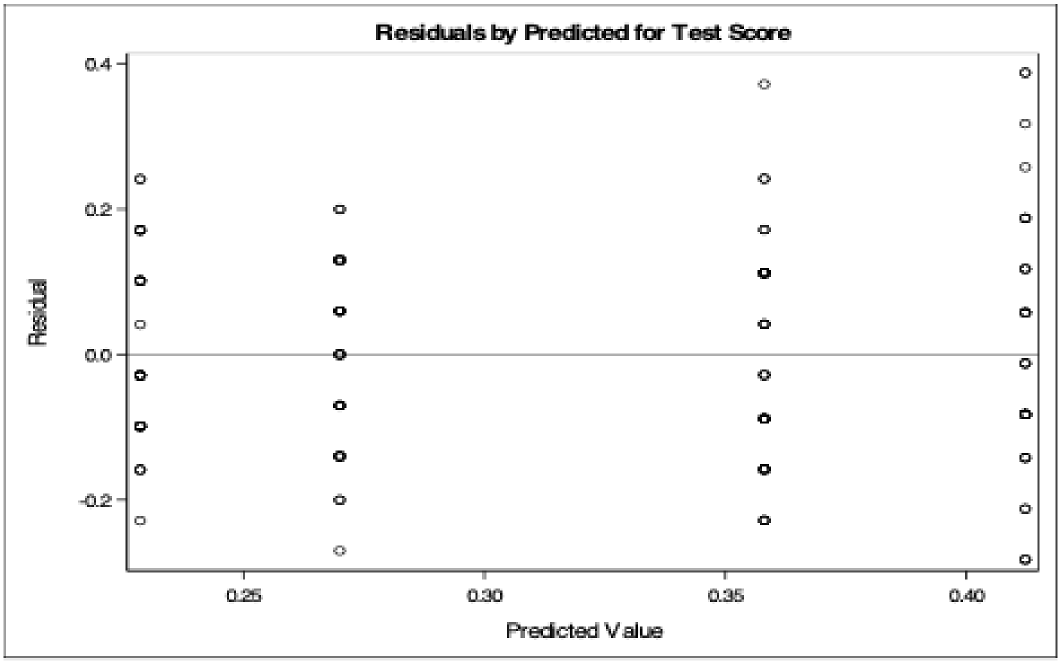
Number of Observations Read	1070
Number of Observations Used	144

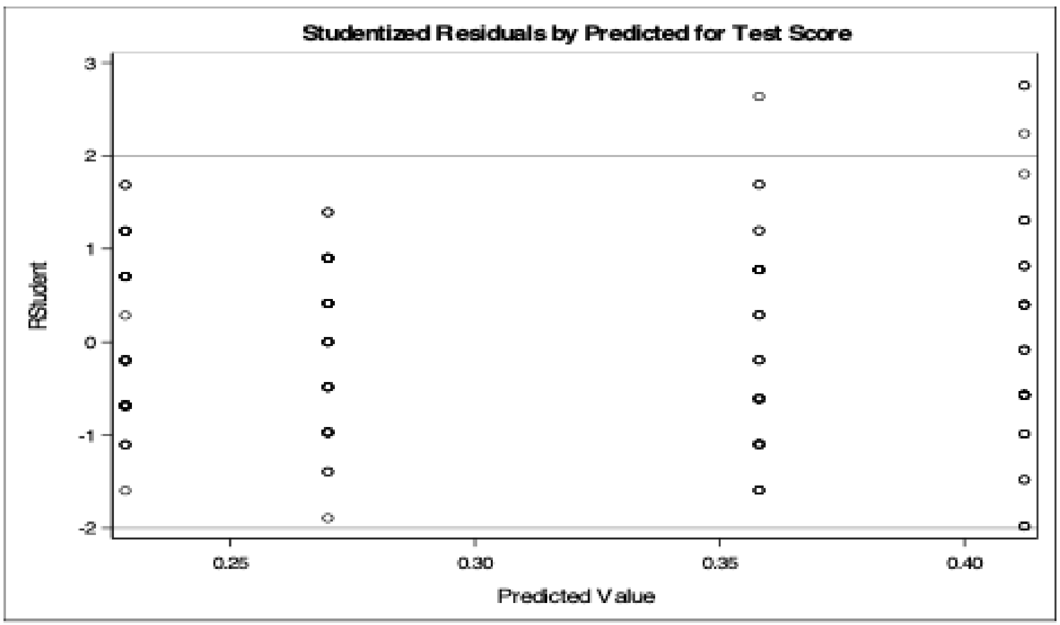
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.74787222	0.24929074	11.68	<.0001
Error	140	2.98781667	0.02134155		
Corrected Total	143	3.73568889			

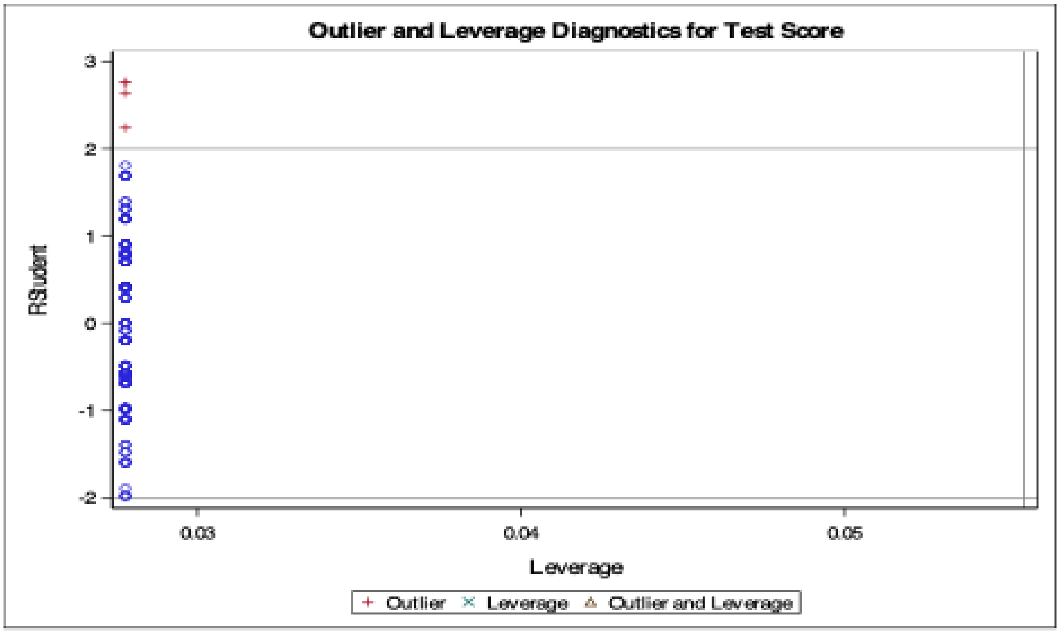
R-Square	Coeff Var	Root MSE	Test Score Mean
0.200197	46.05209	0.146087	0.317222

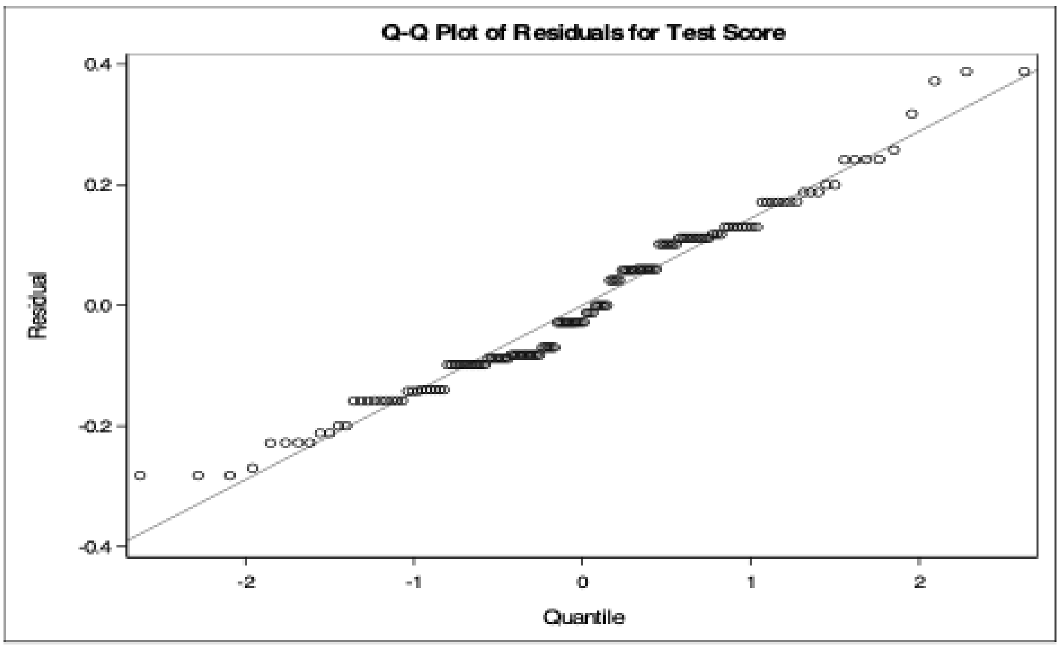
Source	DF	Type I SS	Mean Square	F Value	Pr > F
Group	1	0.08217778	0.08217778	3.85	0.0517
Pre-Post	1	0.66422500	0.66422500	31.12	<.0001
Pre-Post*Group	1	0.00146944	0.00146944	0.07	0.7934

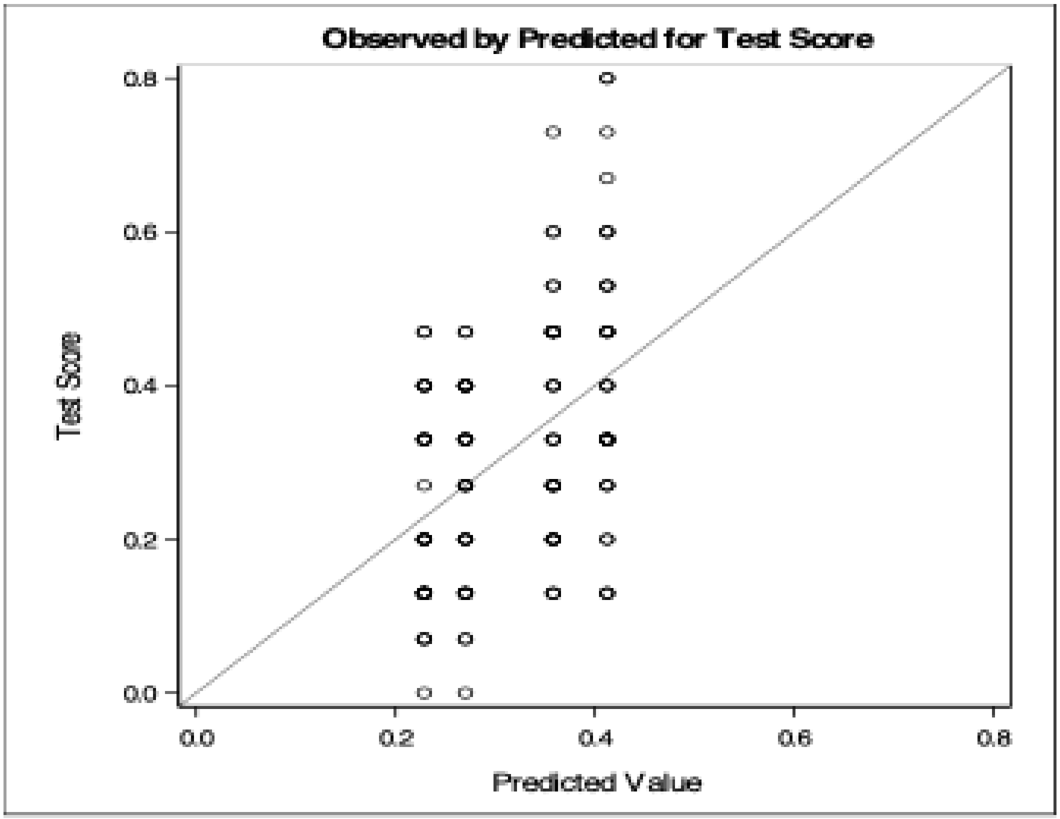
Source	DF	Type III SS	Mean Square	F Value	Pr > F
Group	1	0.03083472	0.03083472	1.44	0.2314
Pre-Post	1	0.66422500	0.66422500	31.12	<.0001
Pre-Post*Group	1	0.00146944	0.00146944	0.07	0.7934

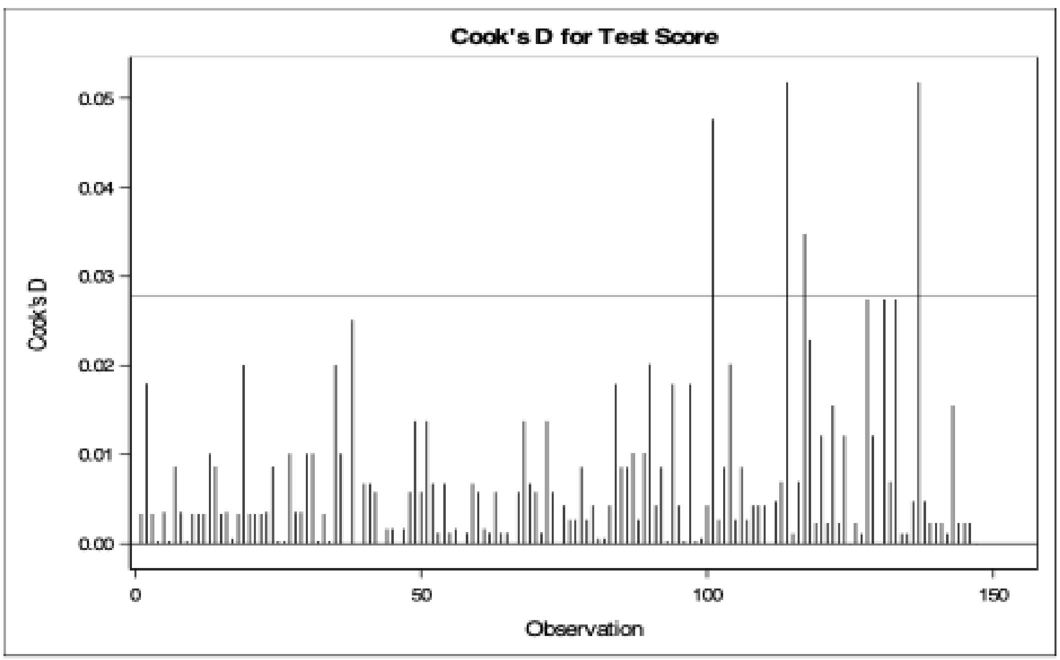

Residuals by Predicted for Test Score
Residual
0.4
0.2
0.0
-0.2
0.25
0.30
0.35
0.40
Predicted Value

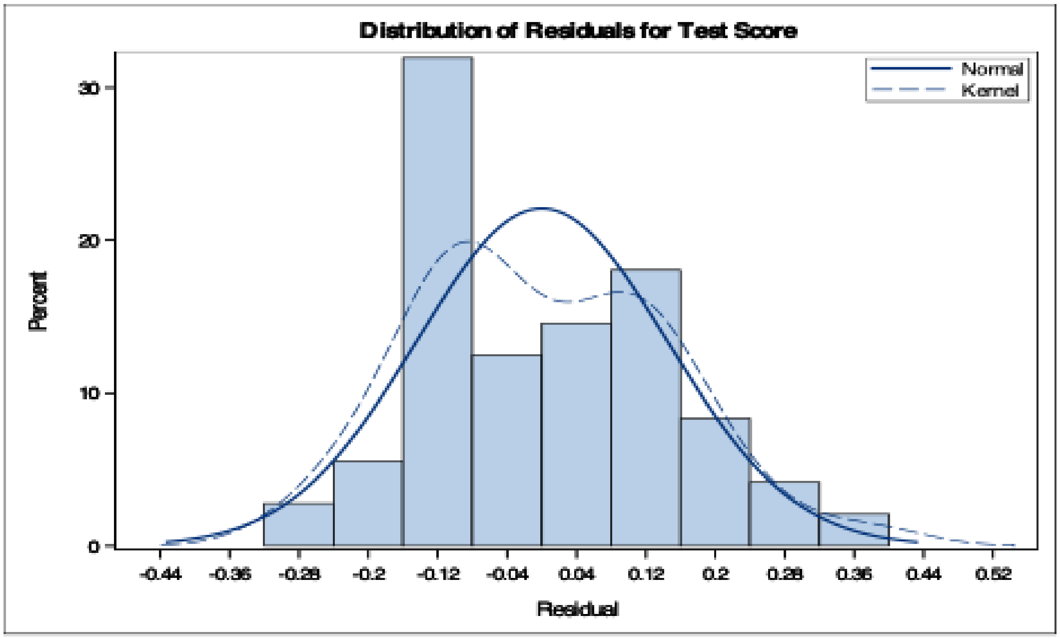

Studentized Residuals by Predicted for Test Score
RStudent
3
2
1
0
-1
-2
0.25
0.30
0.35
0.40
Predicted Value

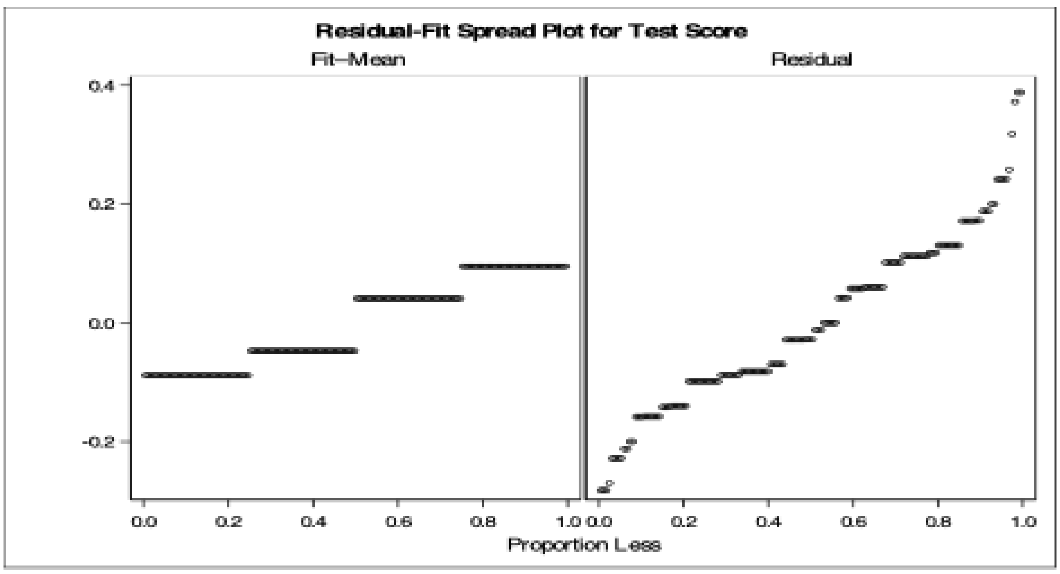

Outlier and Leverage Diagnostics for Test Score
RStudent
3
2
1
0
-1
-2
0.03
0.04
0.05
Leverage
Outlier
Leverage
Outlier and Leverage

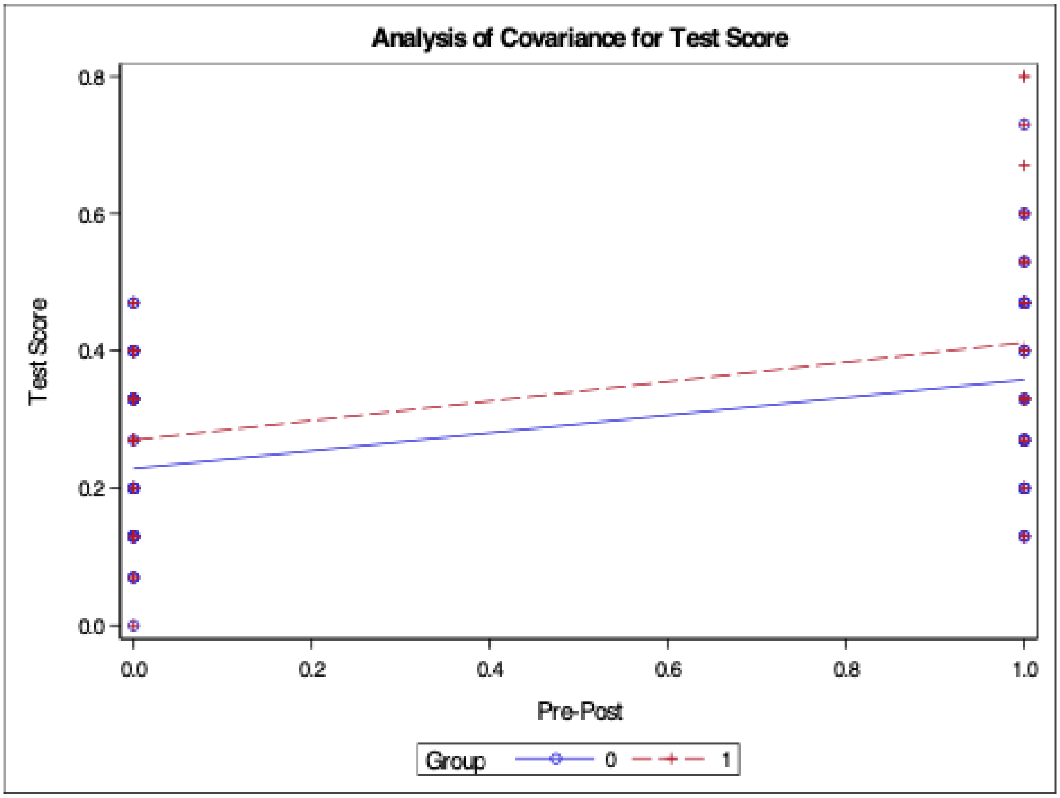

Q-Q Plot of Residuals for Test Score
Residual
0.4
0.2
0.0
-0.2
-0.4
-2
-1
0
1
2
Quantile


Observed by Predicted for Test Score
Test Score
0.0
0.2
0.4
0.6
0.8
Predicted Value

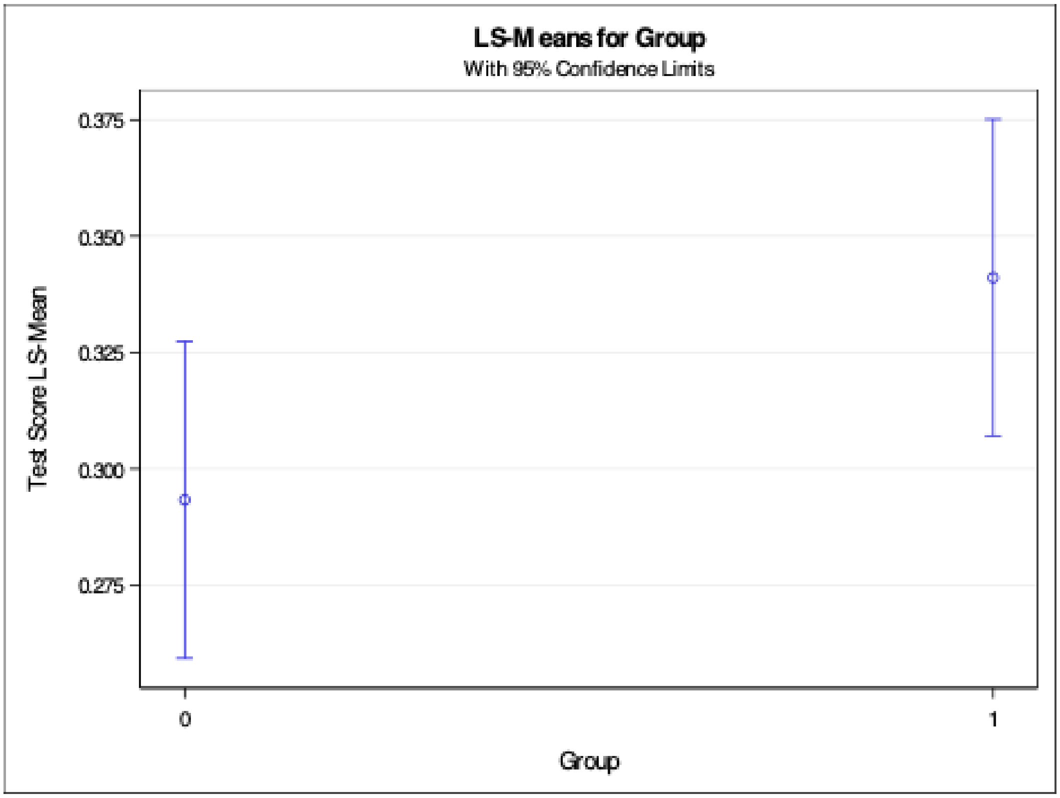

Cook's D for Test Score
Cook's D
0.00
0.01
0.02
0.03
0.04
0.05
0
50
100
150
Observation


Distribution of Residuals for Test Score
Normal
Kernel
Percent
30
20
10
0
-0.44 -0.36 -0.28 -0.2 -0.12 -0.04 0.04 0.12 0.2 0.28 0.36 0.44 0.52
Residual


Residual-Fit Spread Plot for Test Score
Fit–Mean
Residual
0.4
0.2
0.0
-0.2
0.0 0.2 0.4 0.6 0.8 1.0 0.0 0.2 0.4 0.6 0.8 1.0
Proportion Less



Group	Test Score LSMEAN	H0:LSMean1=LSMean2 Pr > |t|
0	0.29333333	0.0517
1	0.34111111	


LS-Means for Group
With 95% Confidence Limits
Test Score LS-Mean
0.375
0.350
0.325
0.300
0.275
0
1
Group

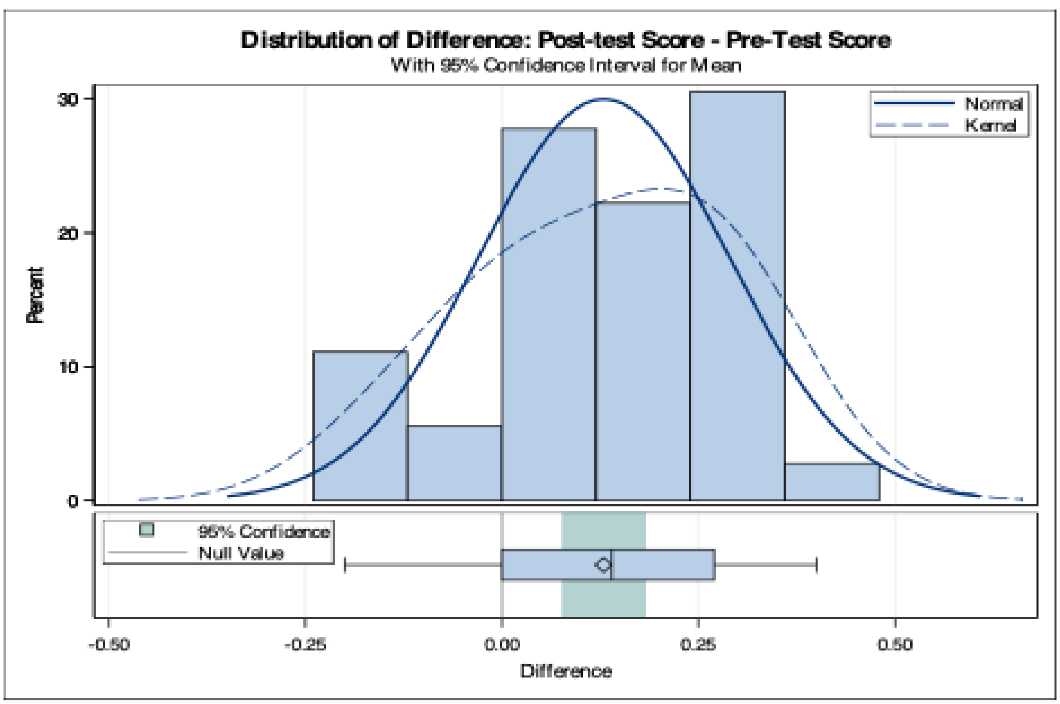
Pre-post Paired T-test Control Group on Teacher-generated assessment

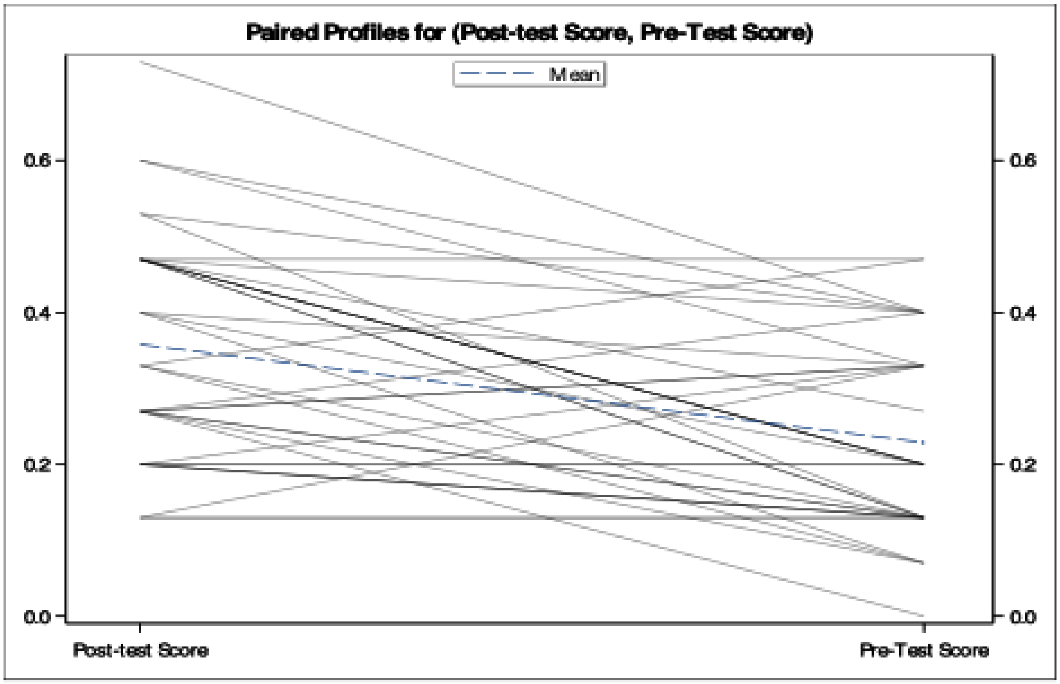
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.953133	**Pr < W**	0.1312
Kolmogorov-Smirnov	**D**	0.142964	**Pr > D**	0.0621
Cramer-von Mises	**W-Sq**	0.103011	**Pr > W-Sq**	0.0990
Anderson-Darling	**A-Sq**	0.623133	**Pr > A-Sq**	0.0974

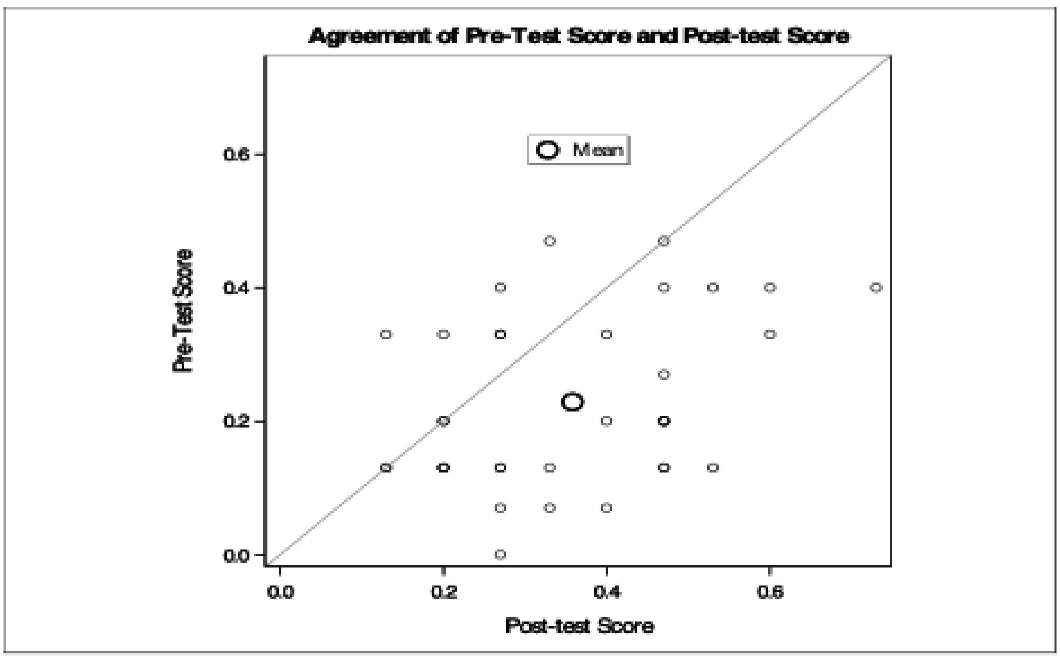
N	Mean	Std Dev	Std Err	Minimum	Maximum
36	0.1294	0.1596	0.0266	-0.2000	0.4000

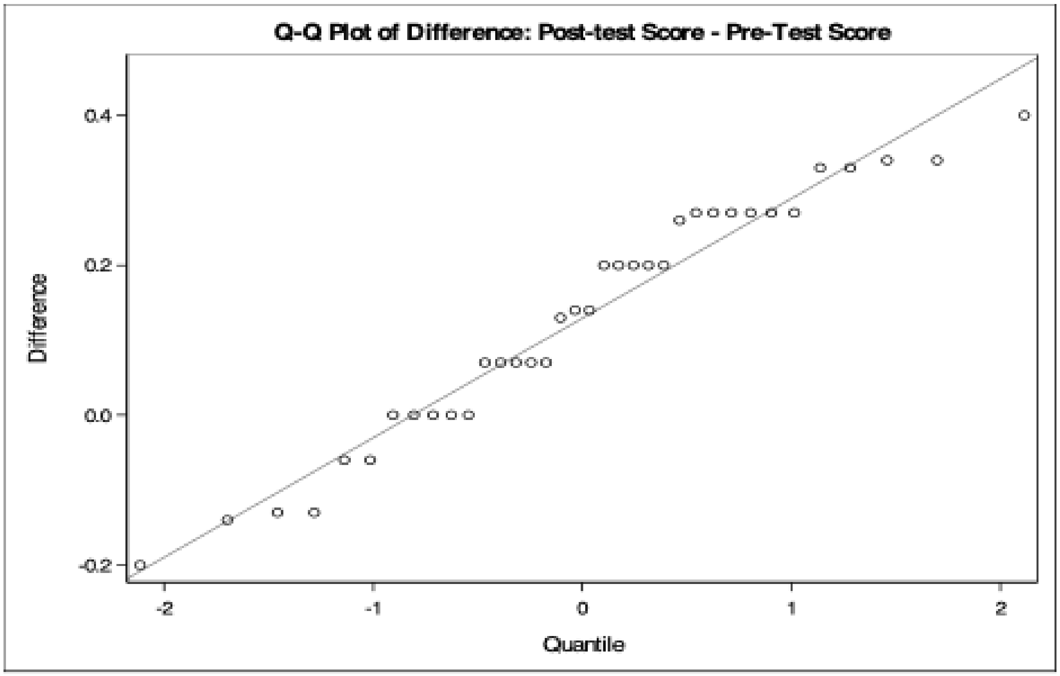
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0.1294	0.0754	0.1835	0.1596	0.1295	0.2082

DF	t Value	Pr > |t|
35	4.87	<.0001




Paired Profiles for (Post-test Score, Pre-Test Score)
Mean
0.0
0.2
0.4
0.6
Post-test Score
Pre-Test Score


Agreement of Pre-Test Score and Post-test Score
Mean
Pre-Test Score
Post-test Score
0.0
0.2
0.4
0.6



Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	4.865069	**Pr > |t|**	<.0001
Sign	**M**	9.5	**Pr >= |M|**	0.0009
Signed Rank	**S**	200	**Pr >= |S|**	<.0001

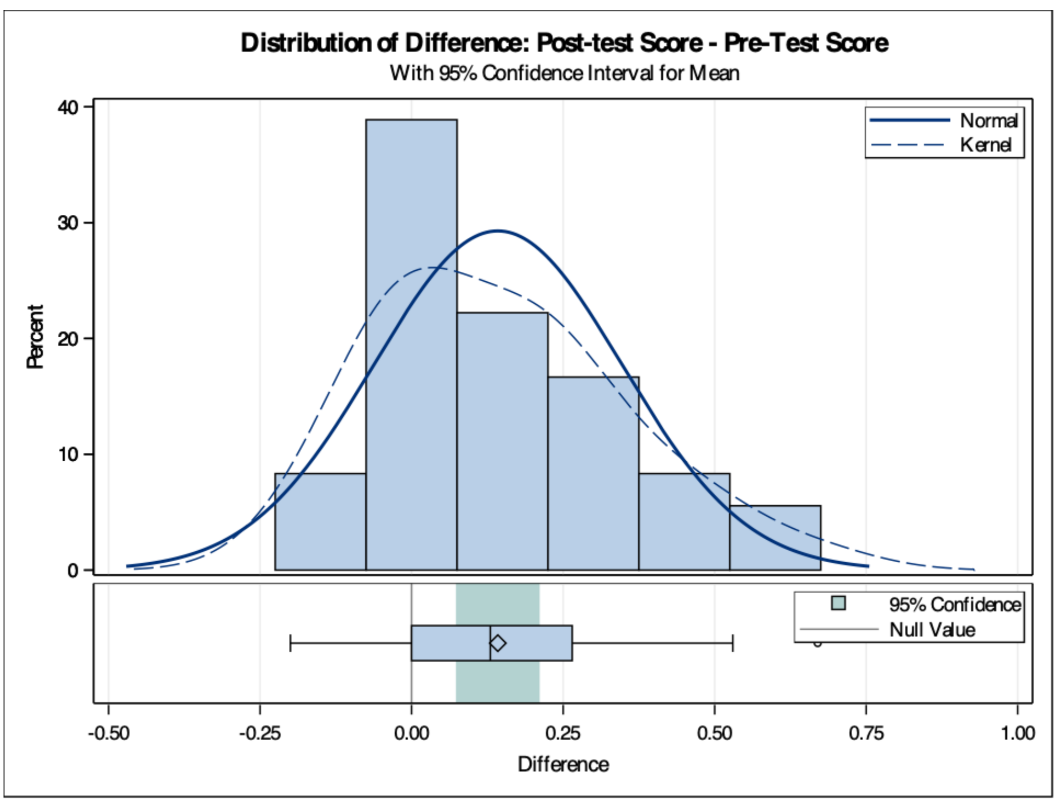
Pre-post Paired T-test Treatment Group on Teacher-generated assessment

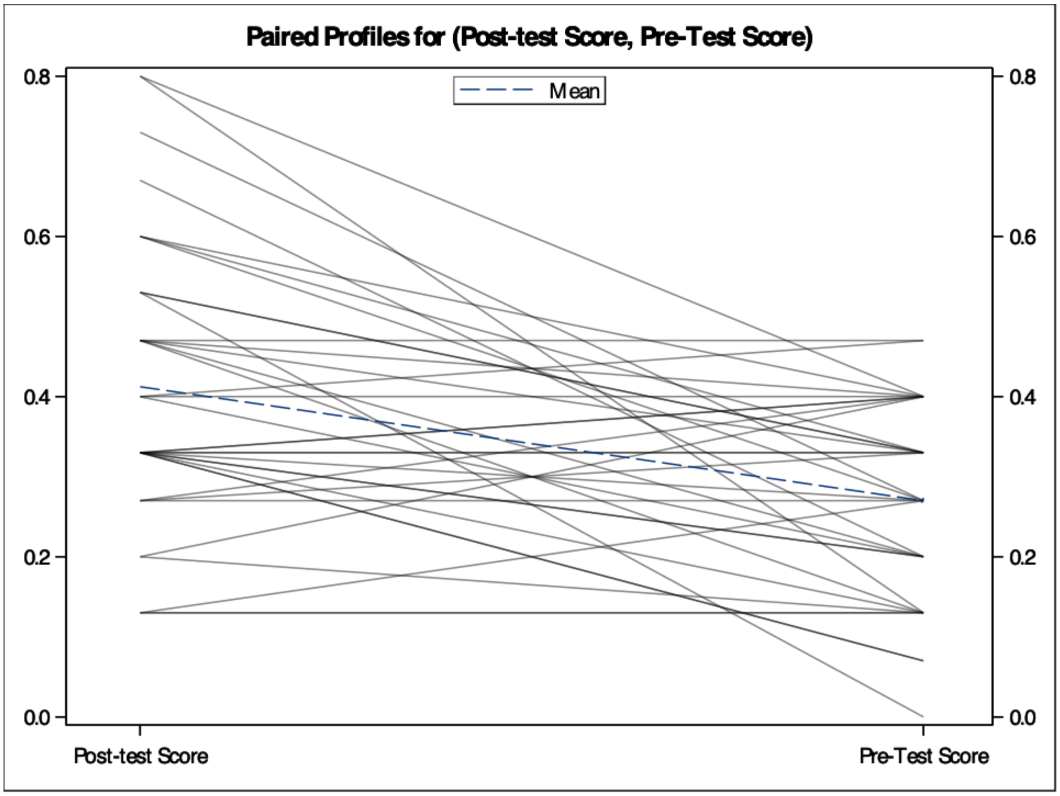
Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.959938	Pr < W	0.2140
Kolmogorov-Smirnov	D	0.145692	Pr > D	0.0511
Cramer-von Mises	W-Sq	0.084357	Pr > W-Sq	0.1822
Anderson-Darling	A-Sq	0.515267	Pr > A-Sq	0.1875

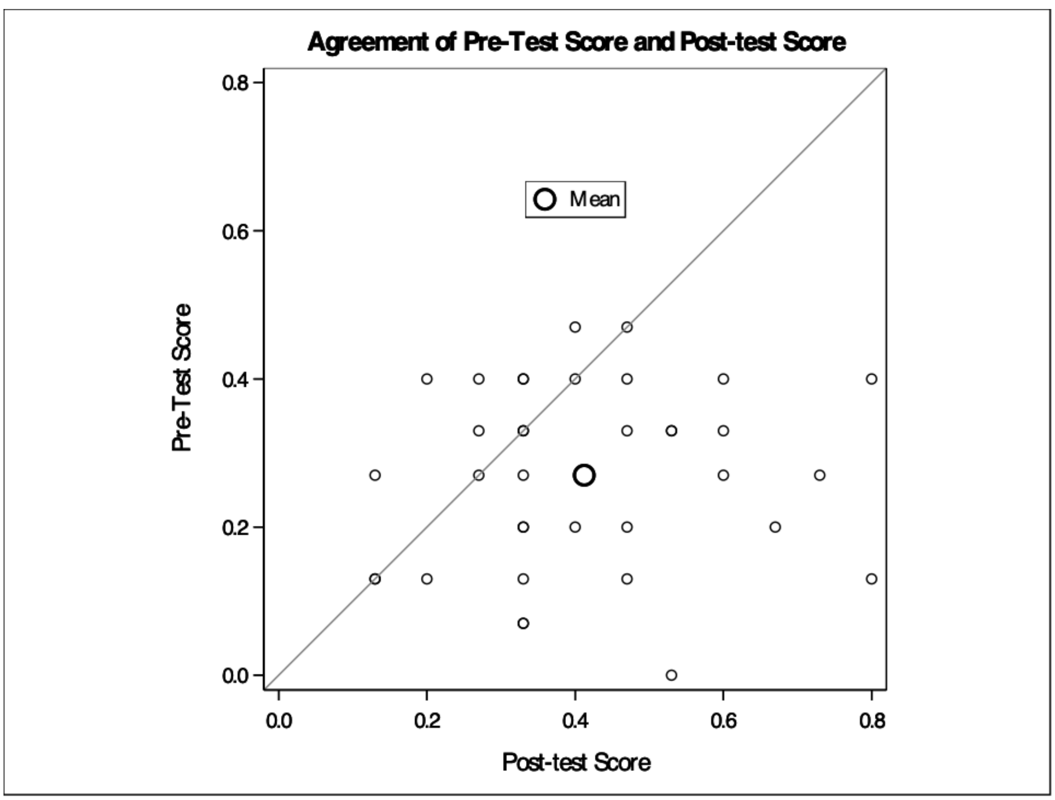
N	Mean	Std Dev	Std Err	Minimum	Maximum
36	0.1422	0.2043	0.0341	-0.2000	0.6700

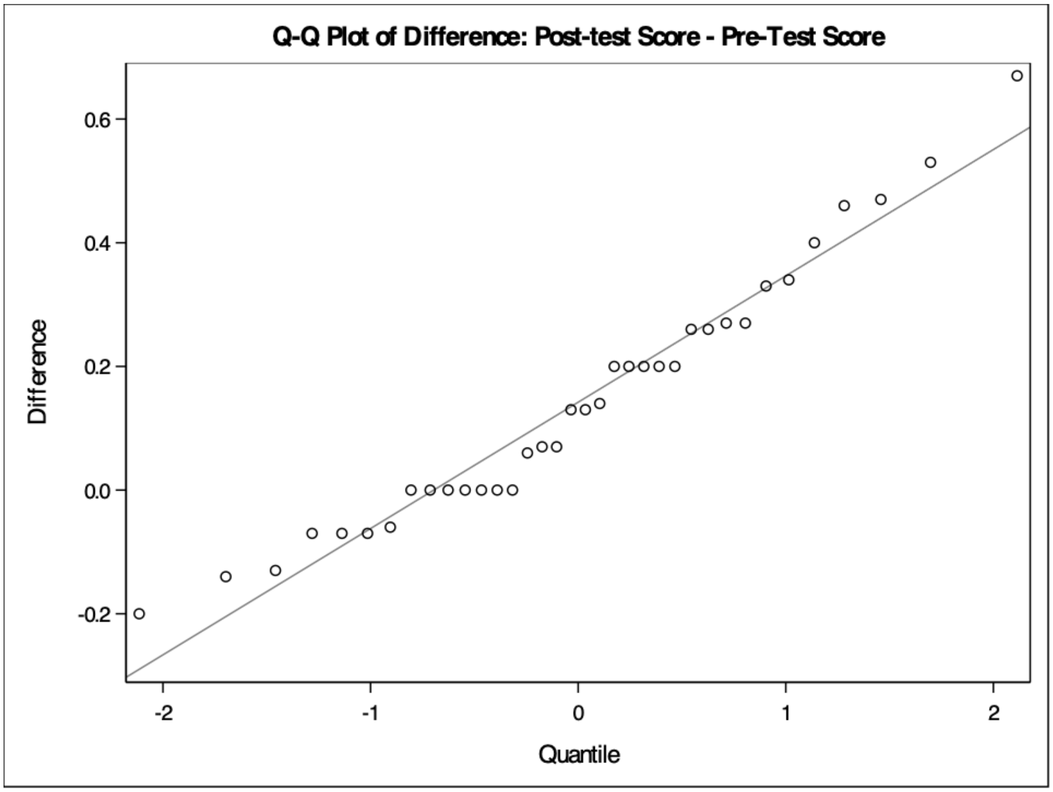
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0.1422	0.0731	0.2114	0.2043	0.1657	0.2665

DF	t Value	Pr > |t|
35	4.18	0.0002




Paired Profiles for (Post-test Score, Pre-Test Score)
Mean
0.8
0.6
0.4
0.2
0.0
Post-test Score
Pre-Test Score


Agreement of Pre-Test Score and Post-test Score
Mean
Pre-Test Score
0.8
0.6
0.4
0.2
0.0
Post-test Score



Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	4.176328	**Pr > |t|**	0.0002
Sign	**M**	7.5	**Pr >= |M|**	0.0081
Signed Rank	**S**	166.5	**Pr >= |S|**	<.0001

Post teacher-generated assessment T-test comparing Treatment and Control Groups

Control Group (0)

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.940656	**Pr < W**	0.0532
Kolmogorov-Smirnov	**D**	0.164902	**Pr > D**	0.0145
Cramer-von Mises	**W-Sq**	0.152349	**Pr > W-Sq**	0.0218
Anderson-Darling	**A-Sq**	0.866443	**Pr > A-Sq**	0.0238

Treatment Group (1)

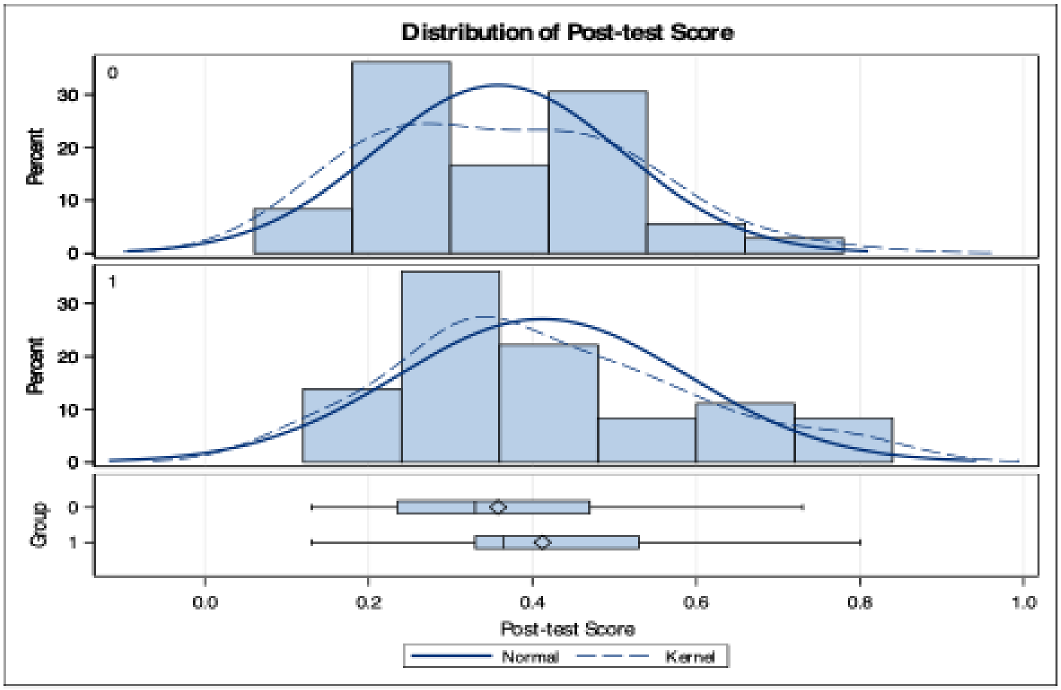
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.948366	**Pr < W**	0.0929
Kolmogorov-Smirnov	**D**	0.179361	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	0.122293	**Pr > W-Sq**	0.0546
Anderson-Darling	**A-Sq**	0.672401	**Pr > A-Sq**	0.0766

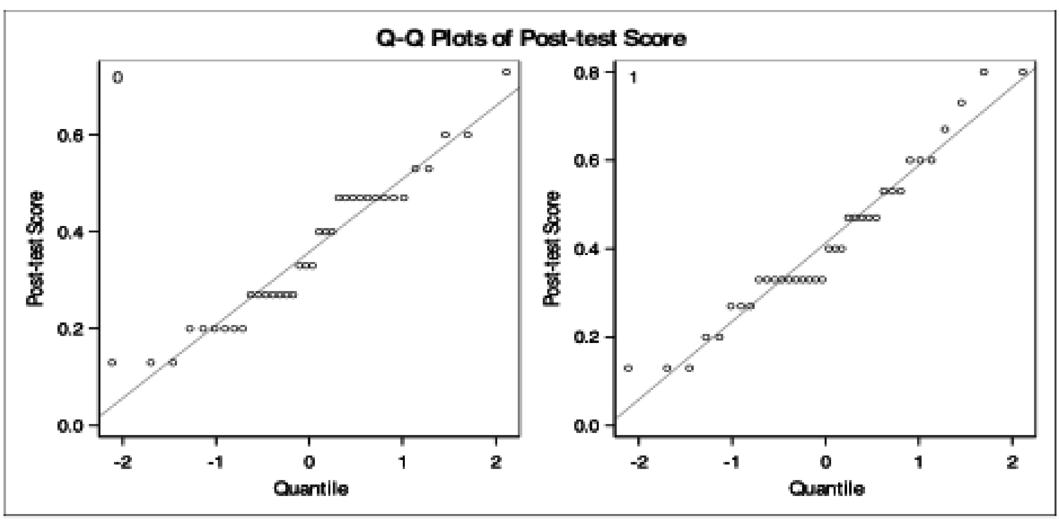
Group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		36	0.3581	0.1507	0.0251	0.1300	0.7300
1		36	0.4122	0.1765	0.0294	0.1300	0.8000
Diff (1-2)	**Pooled**		-0.0542	0.1641	0.0387		
Diff (1-2)	**Satterthwaite**		-0.0542		0.0387		

Group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		0.3581	0.3071	0.4091	0.1507	0.1223	0.1966
1		0.4122	0.3525	0.4719	0.1765	0.1431	0.2302
Diff (1-2)	**Pooled**	-0.0542	-0.1313	0.0230	0.1641	0.1409	0.1966
Diff (1-2)	**Satterthwaite**	-0.0542	-0.1313	0.0230			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	70	-1.40	0.1658
Satterthwaite	Unequal	68.328	-1.40	0.1659

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	35	35	1.37	0.3553

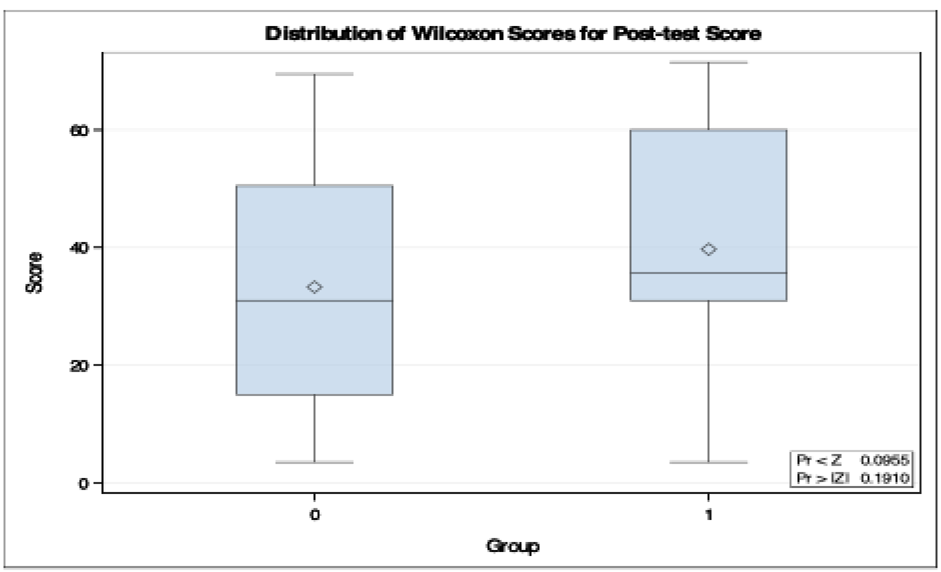




Wilcoxon Scores (Rank Sums) for Variable Post-test Score Classified by Variable Group					
Group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
0	36	1198.50	1314.0	87.945566	33.291667
1	36	1429.50	1314.0	87.945566	39.708333
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	Z	Pr < Z	Pr > |Z|	Pr < Z	Pr > |Z|
1198.500	-1.3076	0.0955	0.1910	0.0976	0.1952
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
1.7248	1	0.1891



Appendix F - Statistical Analysis for ATMI Subscale Self-confidence

Pre-survey T-test comparing Treatment and Control Groups

Control Group (0)

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.903633	**Pr < W**	0.0043
Kolmogorov-Smirnov	**D**	0.129549	**Pr > D**	0.1280
Cramer-von Mises	**W-Sq**	0.103295	**Pr > W-Sq**	0.0983
Anderson-Darling	**A-Sq**	0.790164	**Pr > A-Sq**	0.0384

Treatment Group (1)

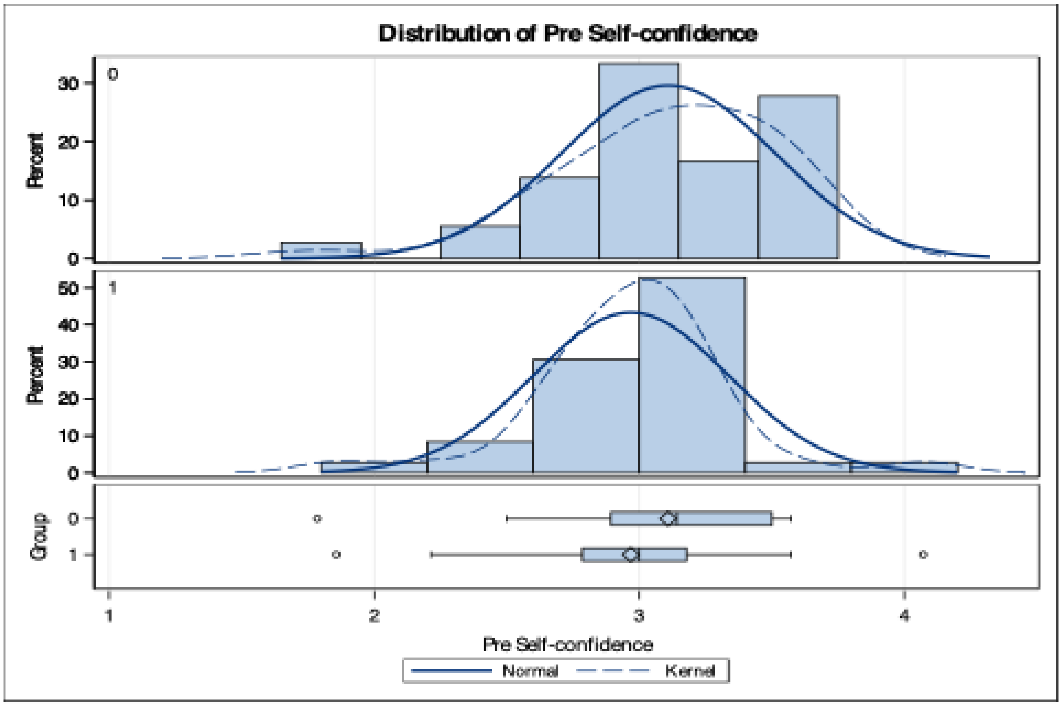
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.92475	**Pr < W**	0.0174
Kolmogorov-Smirnov	**D**	0.138845	**Pr > D**	0.0787
Cramer-von Mises	**W-Sq**	0.140925	**Pr > W-Sq**	0.0308
Anderson-Darling	**A-Sq**	0.92498	**Pr > A-Sq**	0.0183

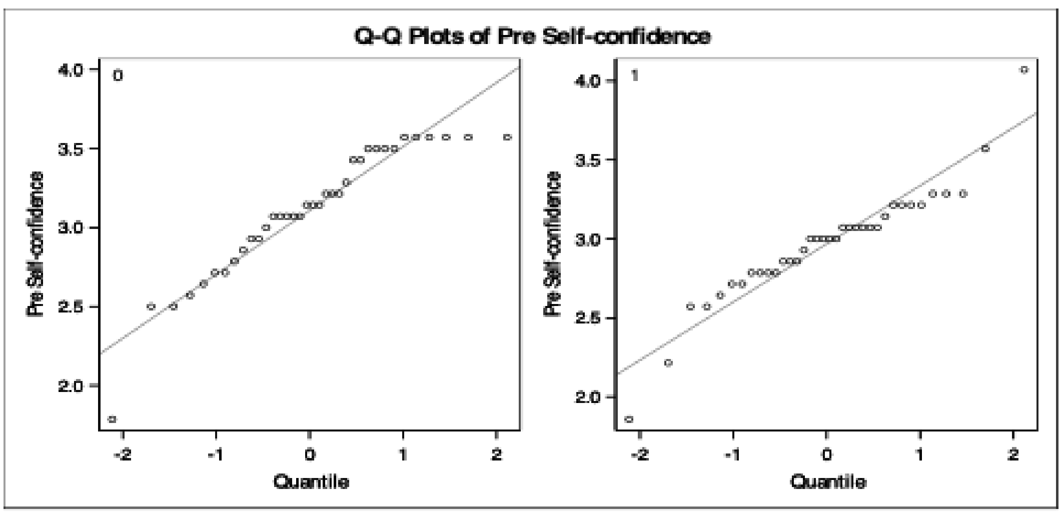
Group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		36	3.1091	0.4046	0.0674	1.7857	3.5714
1		36	2.9683	0.3684	0.0614	1.8571	4.0714
Diff (1-2)	**Pooled**		0.1409	0.3869	0.0912		
Diff (1-2)	**Satterthwaite**		0.1409		0.0912		

Group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3.1091	2.9722	3.2460	0.4046	0.3282	0.5278
1		2.9683	2.8436	3.0929	0.3684	0.2988	0.4805
Diff (1-2)	**Pooled**	0.1409	-0.0410	0.3228	0.3869	0.3321	0.4636
Diff (1-2)	**Satterthwaite**	0.1409	-0.0410	0.3228			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	70	1.54	0.1269
Satterthwaite	Unequal	69.393	1.54	0.1270

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	35	35	1.21	0.5819

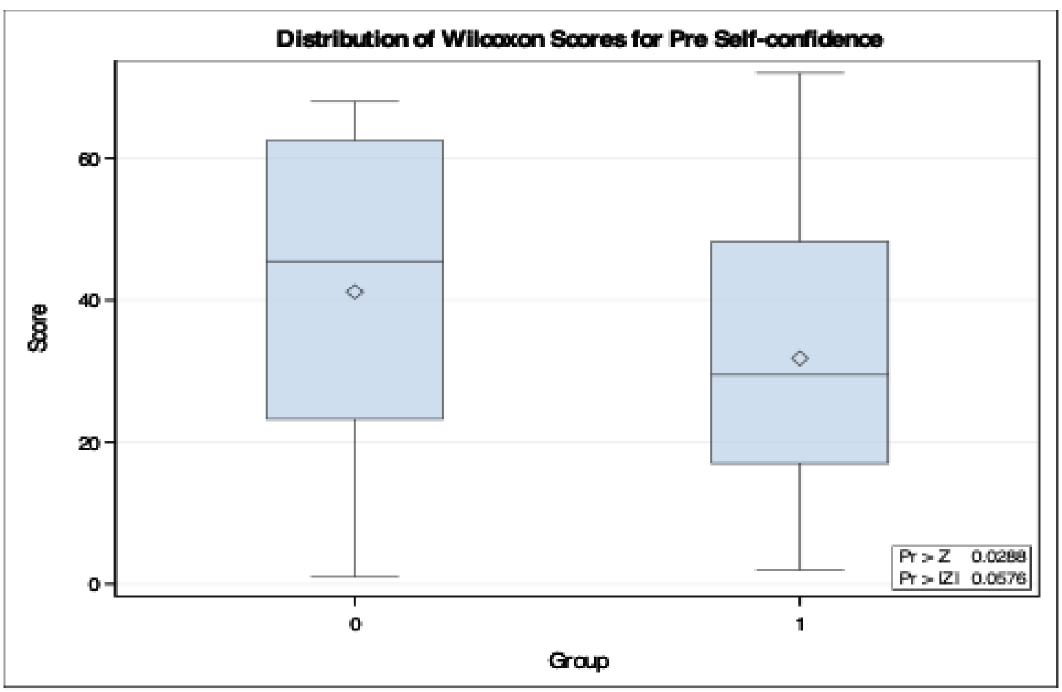




Wilcoxon Scores (Rank Sums) for Variable Pre Self-confidence Classified by Variable Group					
Group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
0	36	1482.50	1314.0	88.471528	41.180556
1	36	1145.50	1314.0	88.471528	31.819444
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	Z	Pr > Z	Pr > |Z|	Pr > Z	Pr > |Z|
1482.500	1.8989	0.0288	0.0576	0.0308	0.0616
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
3.6274	1	0.0568



ANCOVA results comparing Treatment and Control Groups

Class Level Information		
Class	**Levels**	**Values**
Group	2	0 1

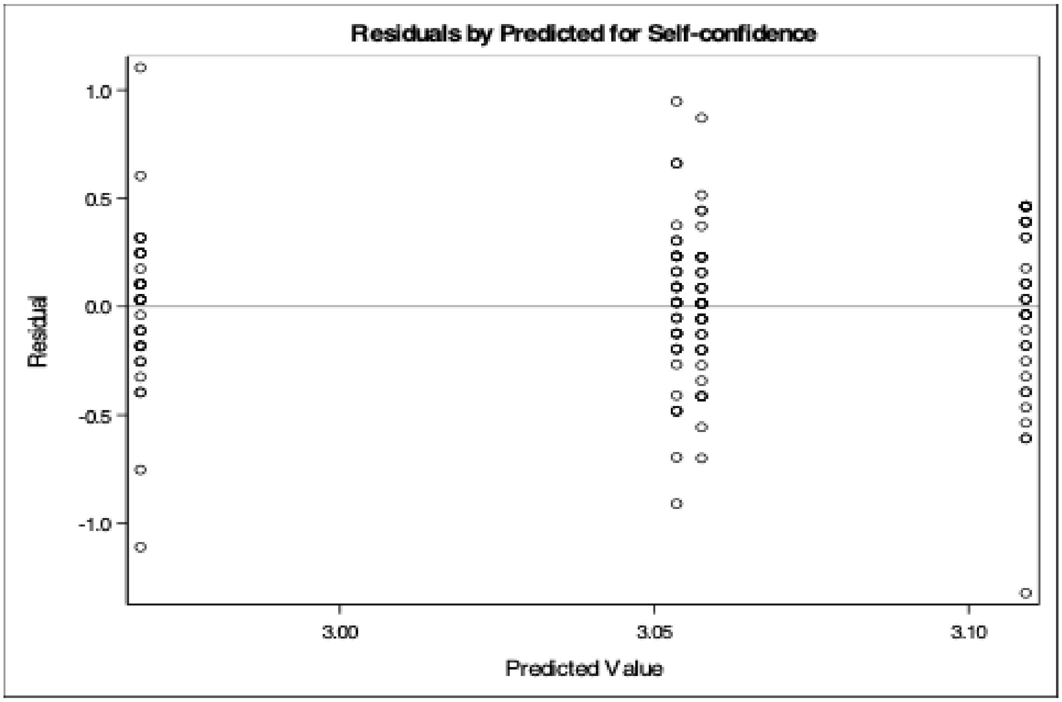
Number of Observations Read	297
Number of Observations Used	144

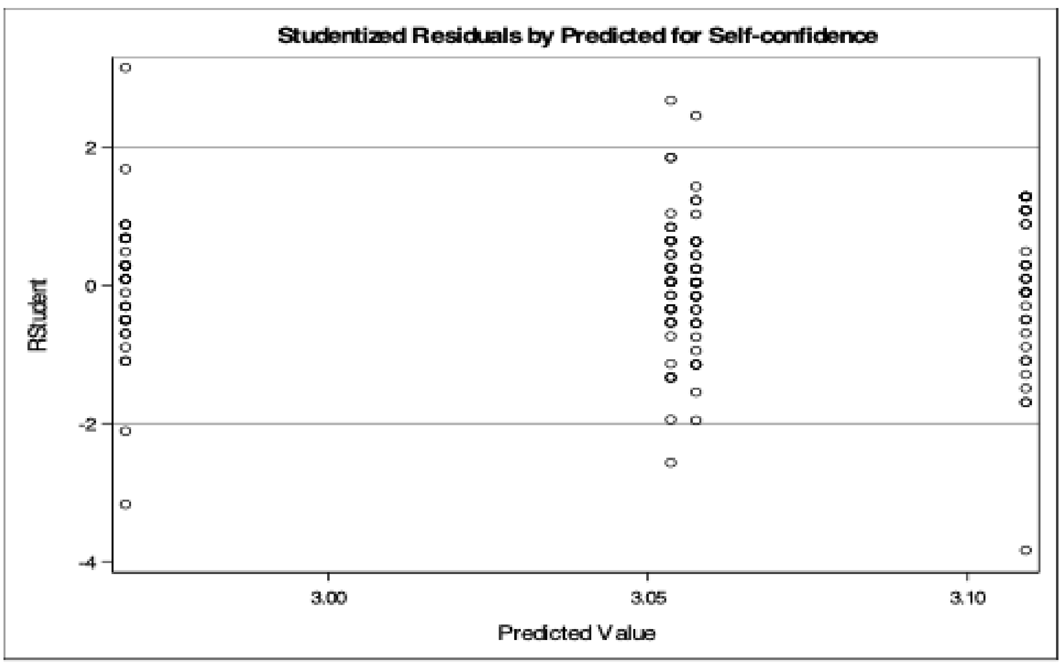
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.36773668	0.12257889	0.91	0.4370
Error	140	18.81760204	0.13441144		
Corrected Total	143	19.18533872			

R-Square	Coeff Var	Root MSE	Self-confidence Mean
0.019168	12.03173	0.366622	3.047123

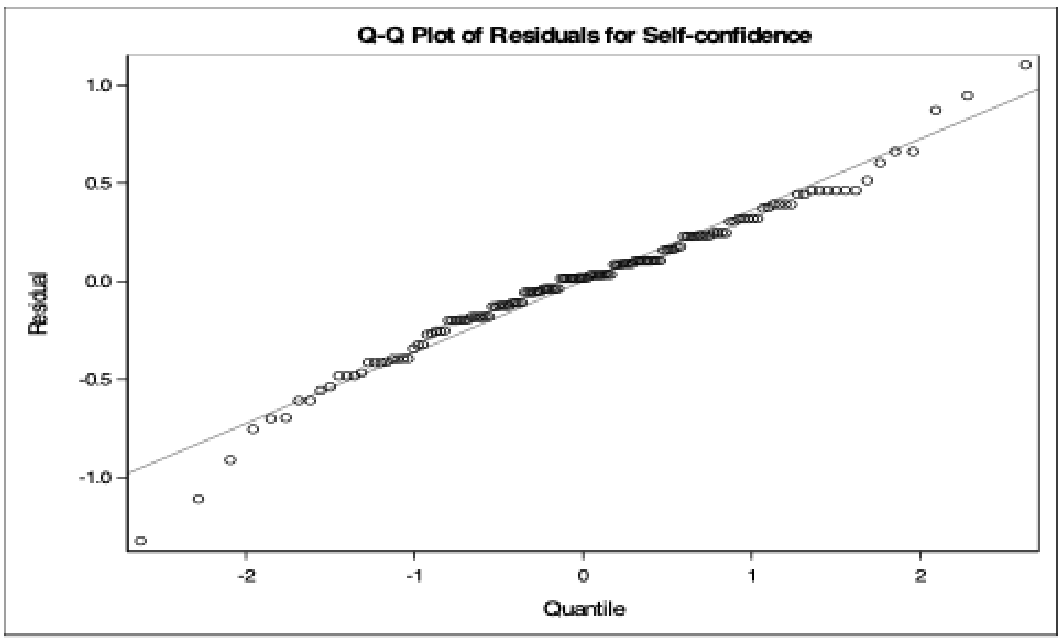
Source	DF	Type I SS	Mean Square	F Value	Pr > F
Group	1	0.18881094	0.18881094	1.40	0.2379
Pre-Post	1	0.01023951	0.01023951	0.08	0.7829
Pre-Post*Group	1	0.16868622	0.16868622	1.25	0.2645

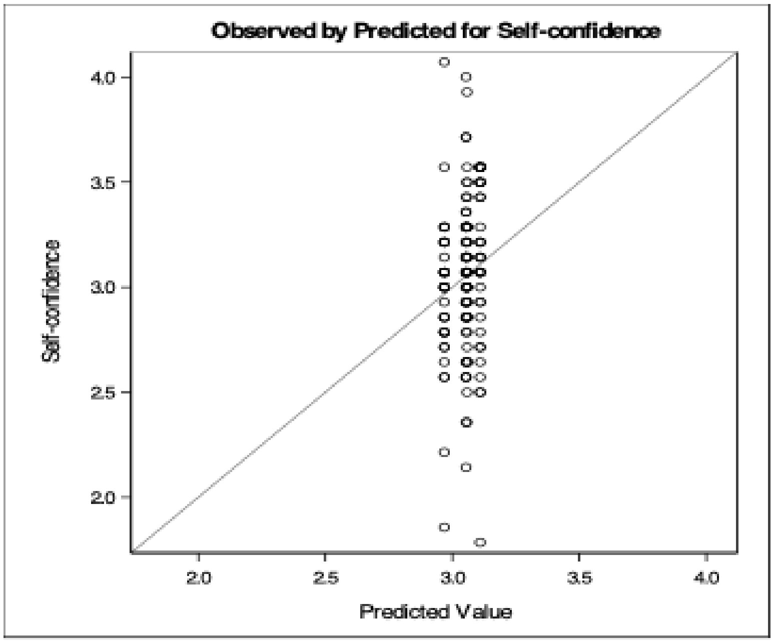
Source	DF	Type III SS	Mean Square	F Value	Pr > F
Group	1	0.35721372	0.35721372	2.66	0.1053
Pre-Post	1	0.01023951	0.01023951	0.08	0.7829
Pre-Post*Group	1	0.16868622	0.16868622	1.25	0.2645

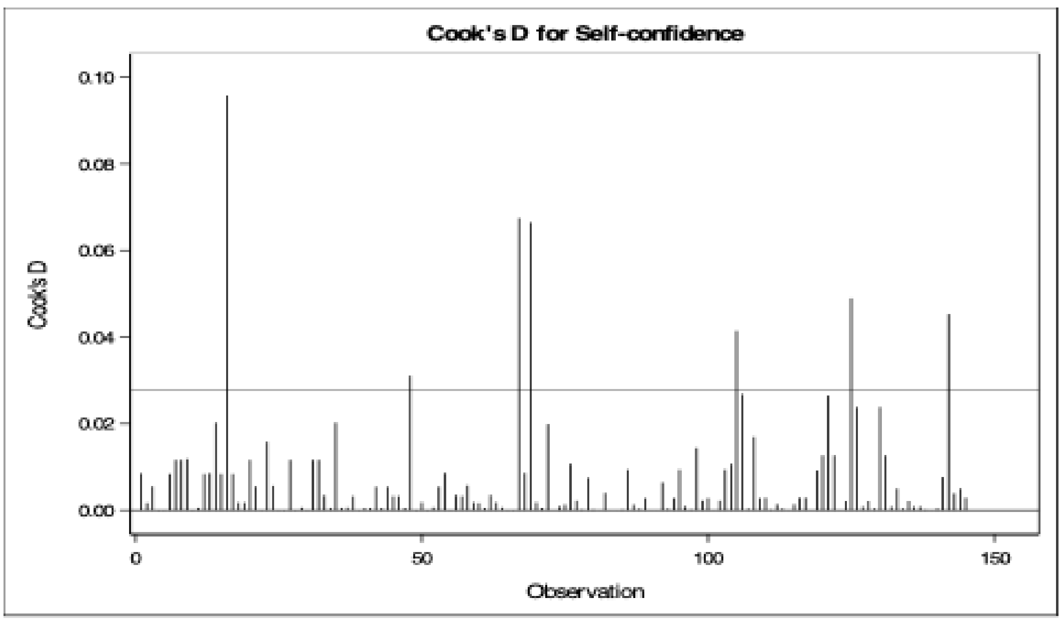

Residuals by Predicted for Self-confidence
Residual
1.0
0.5
0.0
-0.5
-1.0
3.00
3.05
3.10
Predicted Value

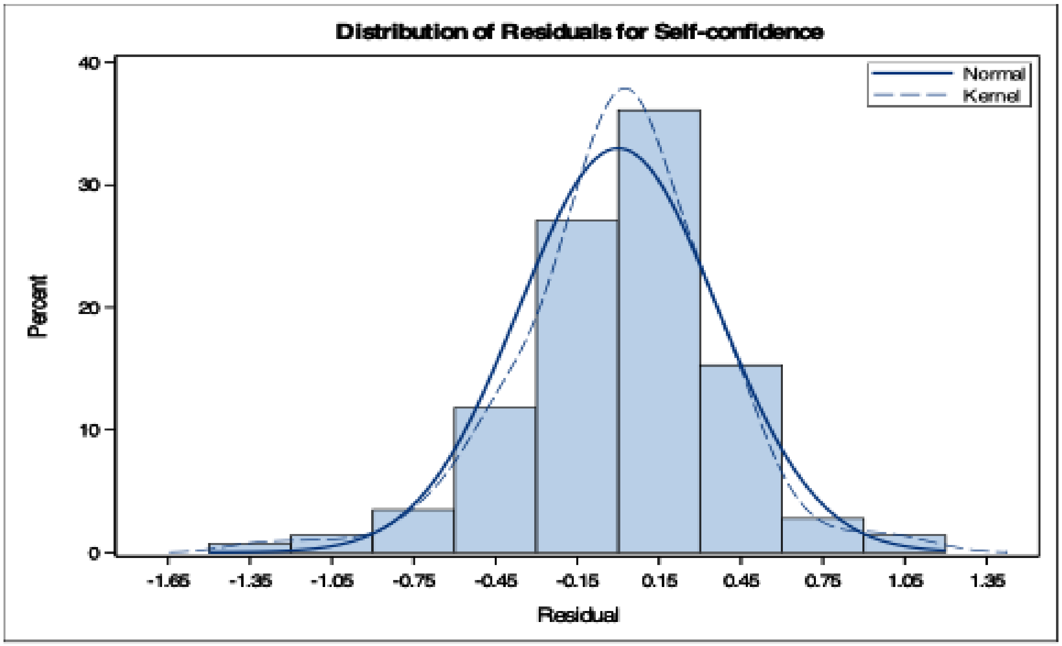

Studentized Residuals by Predicted for Self-confidence
RStudent
2
0
-2
-4
3.00
3.05
3.10
Predicted Value

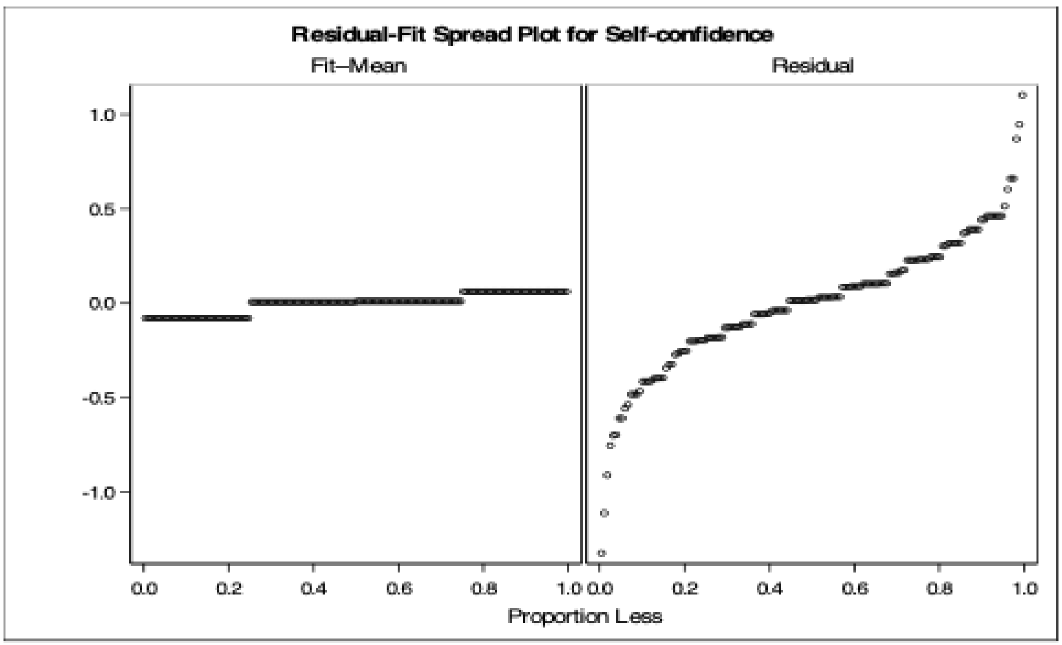

Outlier and Leverage Diagnostics for Self-confidence
RStudent
2
0
-2
-4
0.03
0.04
0.05
Leverage
Outlier
Leverage
Outlier and Leverage

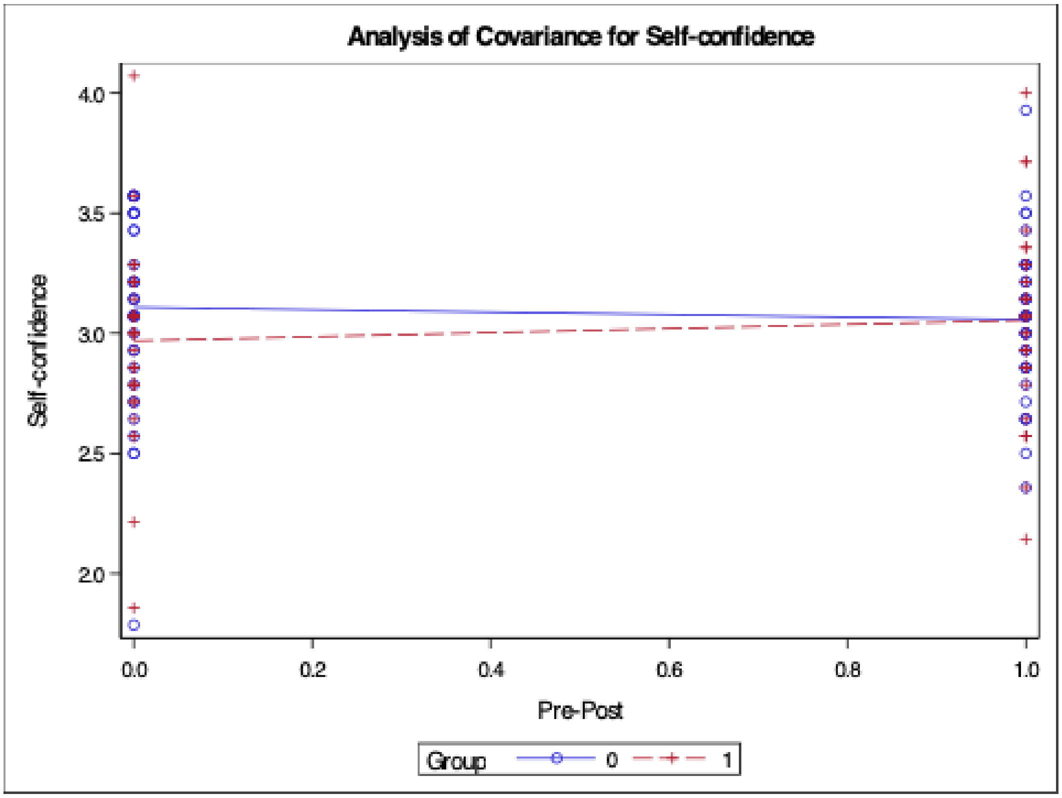

Q-Q Plot of Residuals for Self-confidence
Residual
1.0
0.5
0.0
-0.5
-1.0
-2
-1
0
1
2
Quantile


Observed by Predicted for Self-confidence
Self-confidence
4.0
3.5
3.0
2.5
2.0
2.0
2.5
3.0
3.5
4.0
Predicted Value


Cook's D for Self-confidence
Cook's D
0.10
0.08
0.06
0.04
0.02
0.00
0
50
100
150
Observation


Distribution of Residuals for Self-confidence
Normal
Kernel
Percent
40
30
20
10
0
-1.65
-1.35
-1.05
-0.75
-0.45
-0.15
0.15
0.45
0.75
1.05
1.35
Residual


Residual-Fit Spread Plot for Self-confidence
Fit–Mean
Residual
1.0
0.5
0.0
-0.5
-1.0
0.0
0.2
0.4
0.6
0.8
1.0
Proportion Less



Group	Self-confidence LSMEAN	H0:LSMean1=LSMean2
		Pr > |t|
0	3.08333333	0.2379
1	3.01091270	

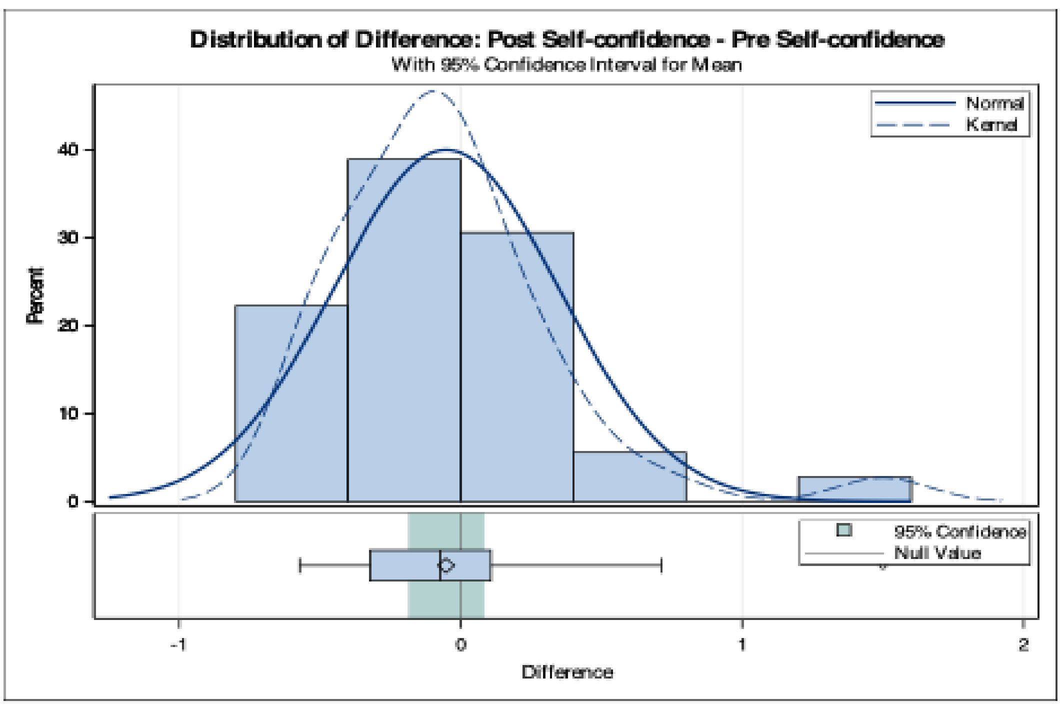
Pre-post Paired T-test Control Group

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.860599	**Pr < W**	0.0003
Kolmogorov-Smirnov	**D**	0.143034	**Pr > D**	0.0618
Cramer-von Mises	**W-Sq**	0.149466	**Pr > W-Sq**	0.0232
Anderson-Darling	**A-Sq**	1.029274	**Pr > A-Sq**	0.0093

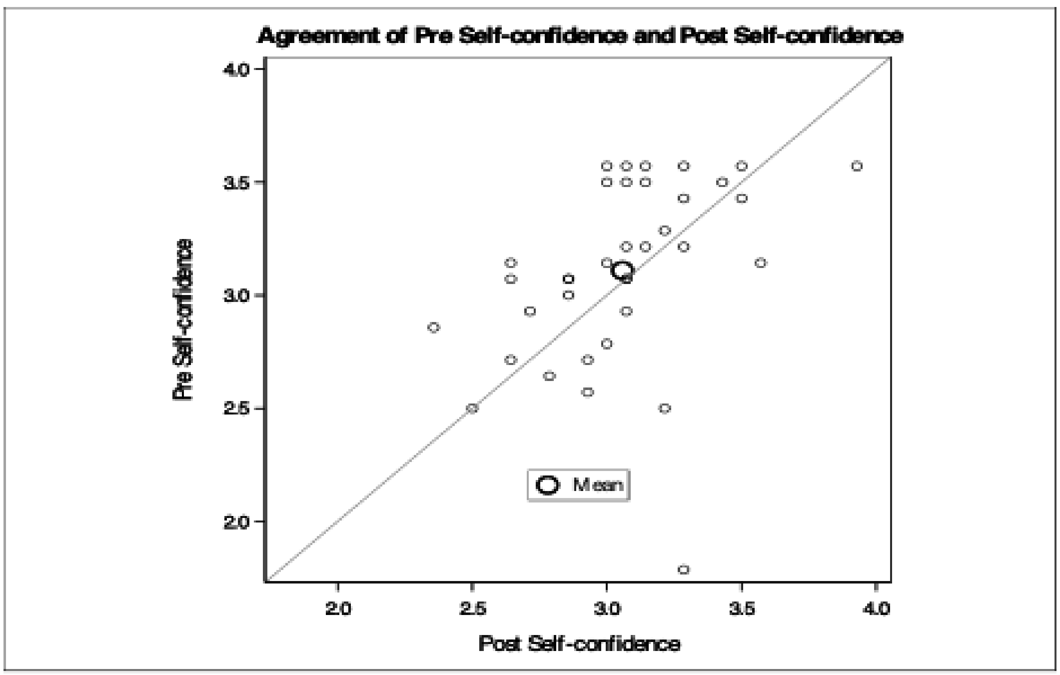
N	Mean	Std Dev	Std Err	Minimum	Maximum
36	-0.0516	0.3992	0.0665	-0.5714	1.5000

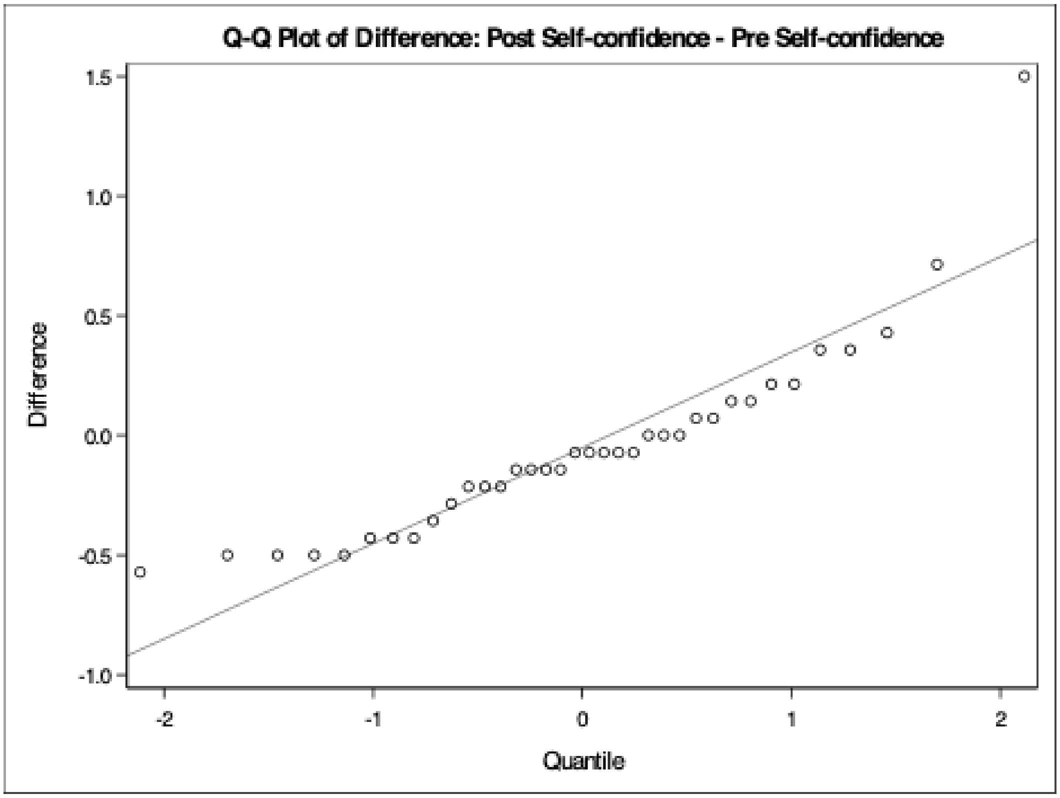
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-0.0516	-0.1867	0.0835	0.3992	0.3238	0.5207

DF	t Value	Pr > |t|
35	-0.78	0.4433




Paired Profiles for (Post Self-confidence, Pre Self-confidence)
Mean
4.0
3.5
3.0
2.5
2.0
Post Self-confidence
Pre Self-confidence


Agreement of Pre Self-confidence and Post Self-confidence
Pre Self-confidence
Mean
2.0
2.5
3.0
3.5
4.0
Post Self-confidence



Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	-0.77536	Pr > |t|	0.4433
Sign	M	-5.5	Pr >= |M|	0.0801
Signed Rank	S	-88.5	Pr >= |S|	0.1147

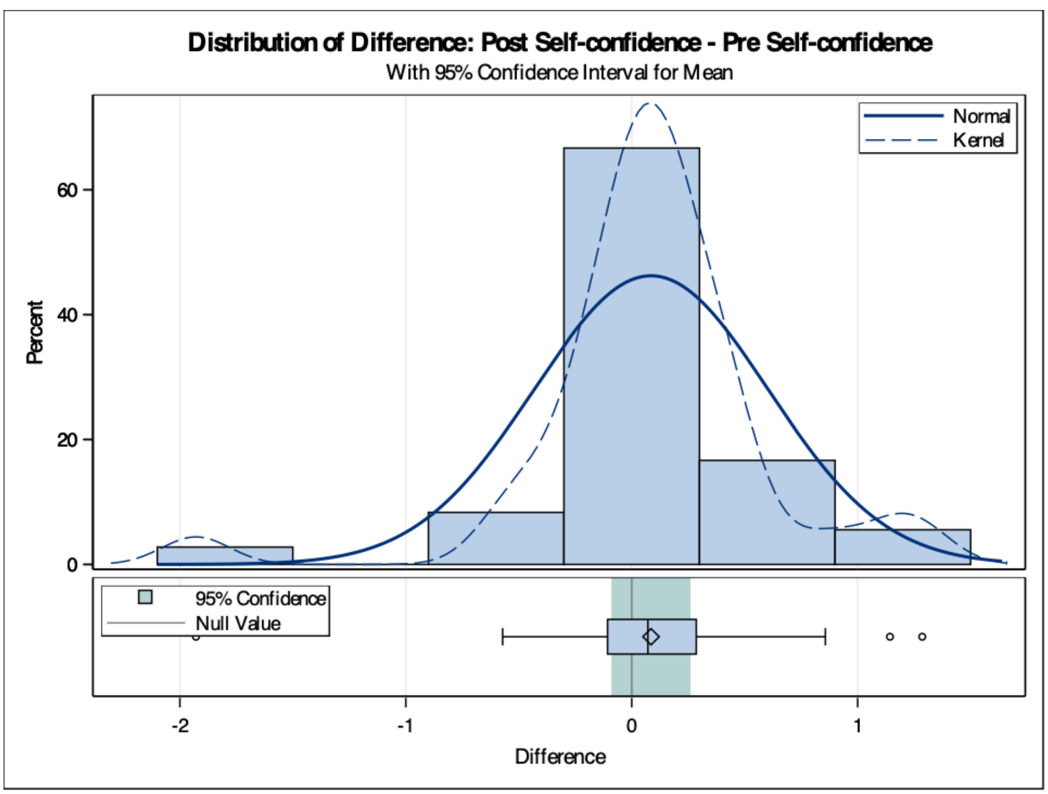
Pre-post Paired T-test Treatment Group

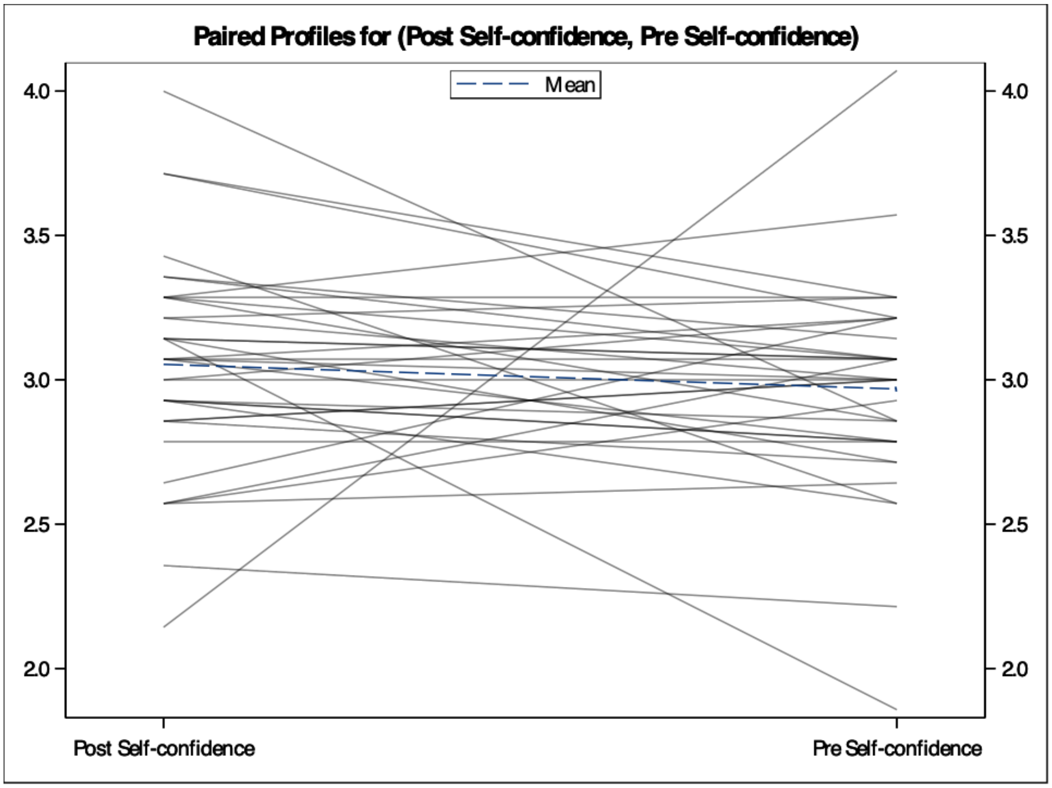
Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.85637	Pr < W	0.0003
Kolmogorov-Smirnov	D	0.163078	Pr > D	0.0166
Cramer-von Mises	W-Sq	0.252586	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	1.512738	Pr > A-Sq	<0.0050

N	Mean	Std Dev	Std Err	Minimum	Maximum
36	0.0853	0.5179	0.0863	-1.9286	1.2857

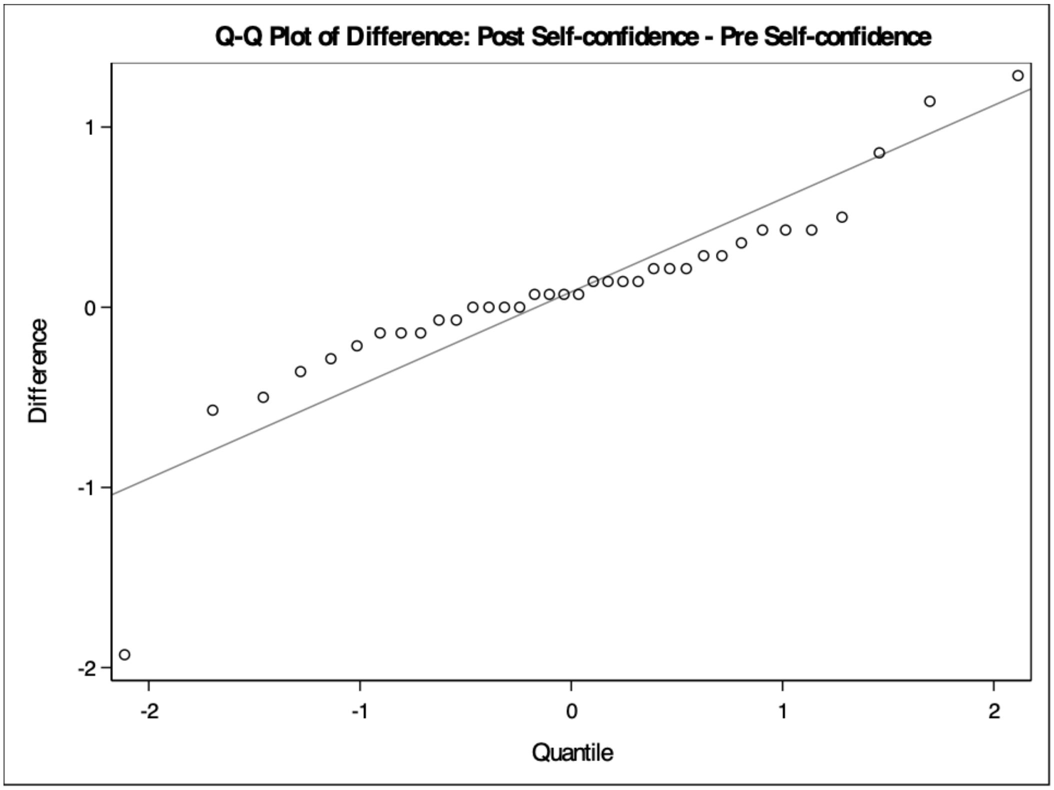
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0.0853	-0.0899	0.2605	0.5179	0.4200	0.6755

DF	t Value	Pr > |t|
35	0.99	0.3297




Paired Profiles for (Post Self-confidence, Pre Self-confidence)
Mean
4.0
3.5
3.0
2.5
2.0
Post Self-confidence
Pre Self-confidence


Agreement of Pre Self-confidence and Post Self-confidence
Mean
Pre Self-confidence
Post Self-confidence
2.0
2.5
3.0
3.5
4.0



Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	**0.988519**	**Pr > |t|**	**0.3297**
Sign	**M**	**5**	**Pr >= |M|**	**0.1102**
Signed Rank	**S**	**87**	**Pr >= |S|**	**0.1040**

Post-survey T-test comparing Treatment and Control Groups

Control Group (0)

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.980559	**Pr < W**	0.7640
Kolmogorov-Smirnov	**D**	0.096661	**Pr > D**	>0.1500
Cramer-von Mises	**W-Sq**	0.056967	**Pr > W-Sq**	>0.2500
Anderson-Darling	**A-Sq**	0.319982	**Pr > A-Sq**	>0.2500

Treatment Group (1)

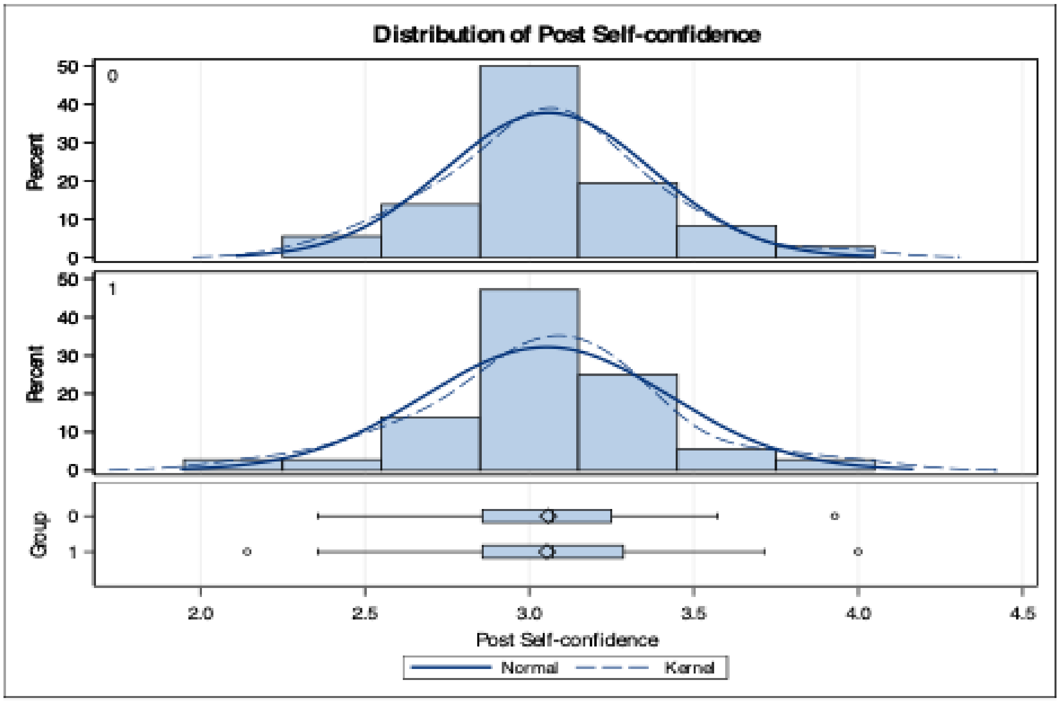
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.974993	**Pr < W**	0.5765
Kolmogorov-Smirnov	**D**	0.104046	**Pr > D**	>0.1500
Cramer-von Mises	**W-Sq**	0.068925	**Pr > W-Sq**	>0.2500
Anderson-Darling	**A-Sq**	0.428224	**Pr > A-Sq**	>0.2500

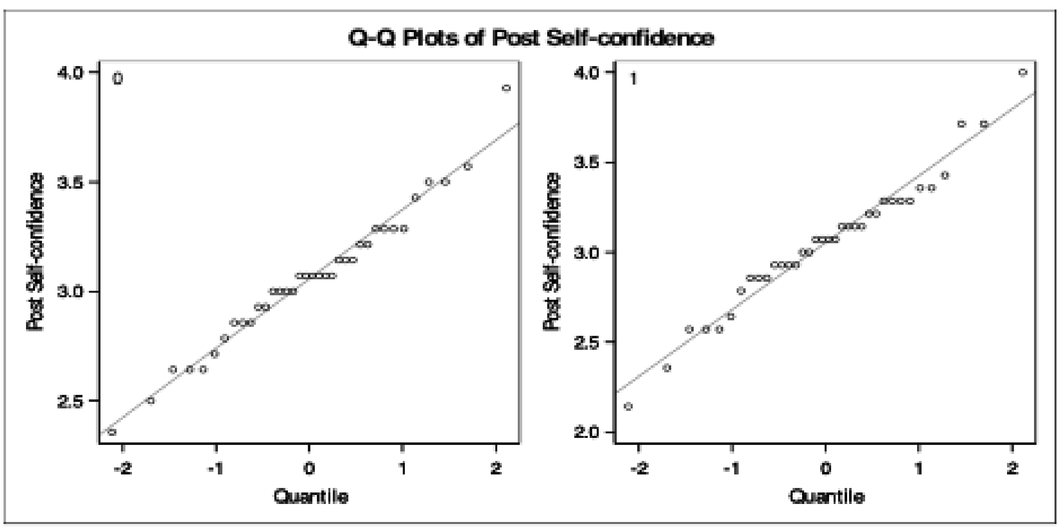
Group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		36	3.0575	0.3166	0.0528	2.3571	3.9286
1		36	3.0536	0.3715	0.0619	2.1429	4.0000
Diff (1-2)	**Pooled**		0.00397	0.3451	0.0814		
Diff (1-2)	**Satterthwaite**		0.00397		0.0814		

Group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3.0575	2.9504	3.1647	0.3166	0.2568	0.4130
1		3.0536	2.9279	3.1793	0.3715	0.3013	0.4846
Diff (1-2)	**Pooled**	0.00397	-0.1583	0.1662	0.3451	0.2962	0.4136
Diff (1-2)	**Satterthwaite**	0.00397	-0.1584	0.1663			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	70	0.05	0.9612
Satterthwaite	Unequal	68.284	0.05	0.9612

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	35	35	1.38	0.3487

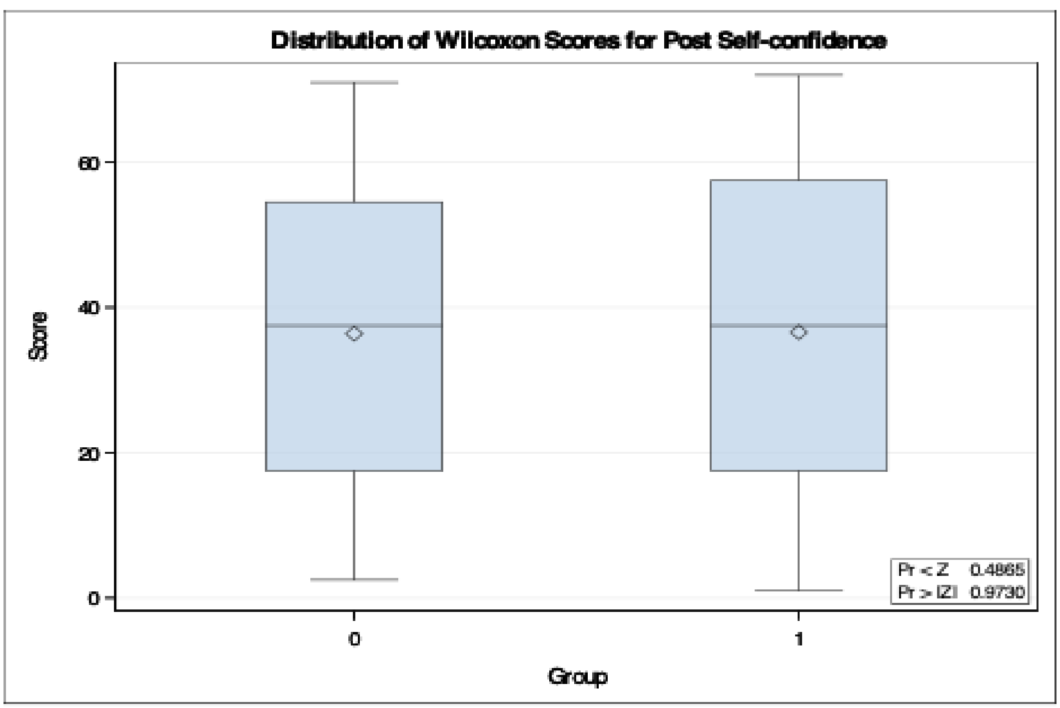




Wilcoxon Scores (Rank Sums) for Variable Post Self-confidence Classified by Variable Group					
Group	**N**	**Sum of Scores**	**Expected Under H0**	**Std Dev Under H0**	**Mean Score**
0	36	1310.50	1314.0	88.477259	36.402778
1	36	1317.50	1314.0	88.477259	36.597222
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	**Z**	**Pr < Z**	**Pr > |Z|**	**Pr < Z**	**Pr > |Z|**
1310.500	-0.0339	0.4865	0.9730	0.4865	0.9730
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	**DF**	**Pr > ChiSq**
0.0016	1	0.9684



Appendix G - Statistical Analysis for ATMI Subscale Value

Pre-survey T-test comparing Treatment and Control Groups

Control Group (0)

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.93761	Pr < W	0.0428
Kolmogorov-Smirnov	D	0.122044	Pr > D	>0.1500
Cramer-von Mises	W-Sq	0.085737	Pr > W-Sq	0.1740
Anderson-Darling	A-Sq	0.620068	Pr > A-Sq	0.0987

Treatment Group (1)

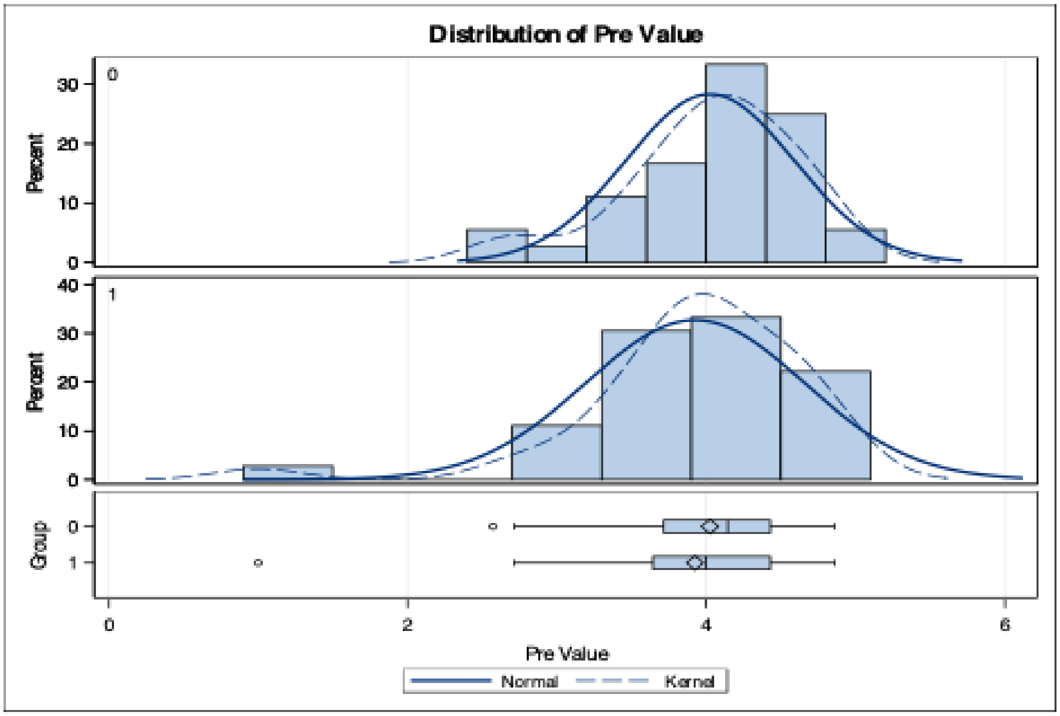
Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.852431	Pr < W	0.0002
Kolmogorov-Smirnov	D	0.157691	Pr > D	0.0228
Cramer-von Mises	W-Sq	0.154389	Pr > W-Sq	0.0207
Anderson-Darling	A-Sq	1.050015	Pr > A-Sq	0.0084

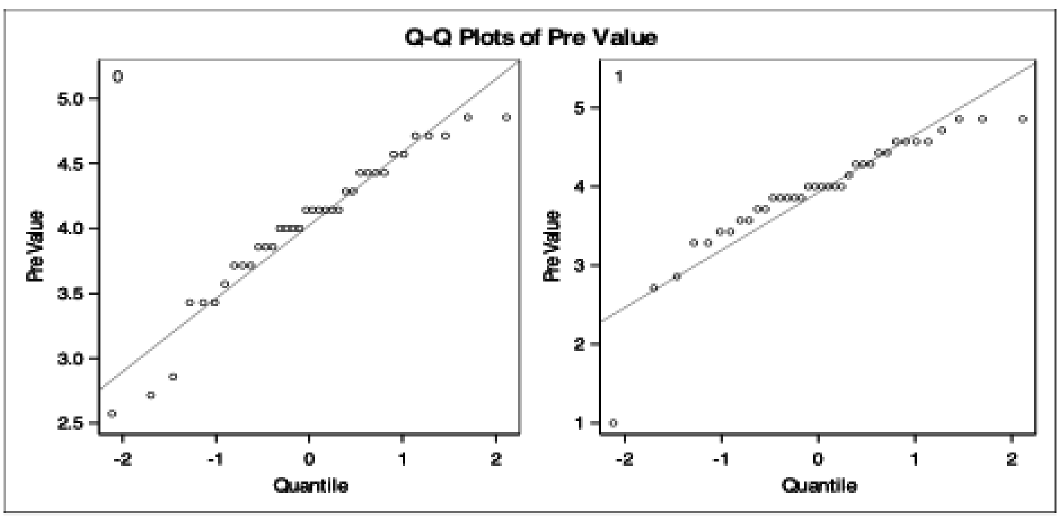
Group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		36	4.0238	0.5637	0.0940	2.5714	4.8571
1		36	3.9246	0.7312	0.1219	1.0000	4.8571
Diff (1-2)	Pooled		0.0992	0.6529	0.1539		
Diff (1-2)	Satterthwaite		0.0992		0.1539		

Group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		4.0238	3.8331	4.2145	0.5637	0.4572	0.7353
1		3.9246	3.6772	4.1720	0.7312	0.5931	0.9538
Diff (1-2)	Pooled	0.0992	-0.2077	0.4061	0.6529	0.5603	0.7823
Diff (1-2)	Satterthwaite	0.0992	-0.2081	0.4065			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	70	0.64	0.5212
Satterthwaite	Unequal	65.744	0.64	0.5214

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	35	35	1.68	0.1286

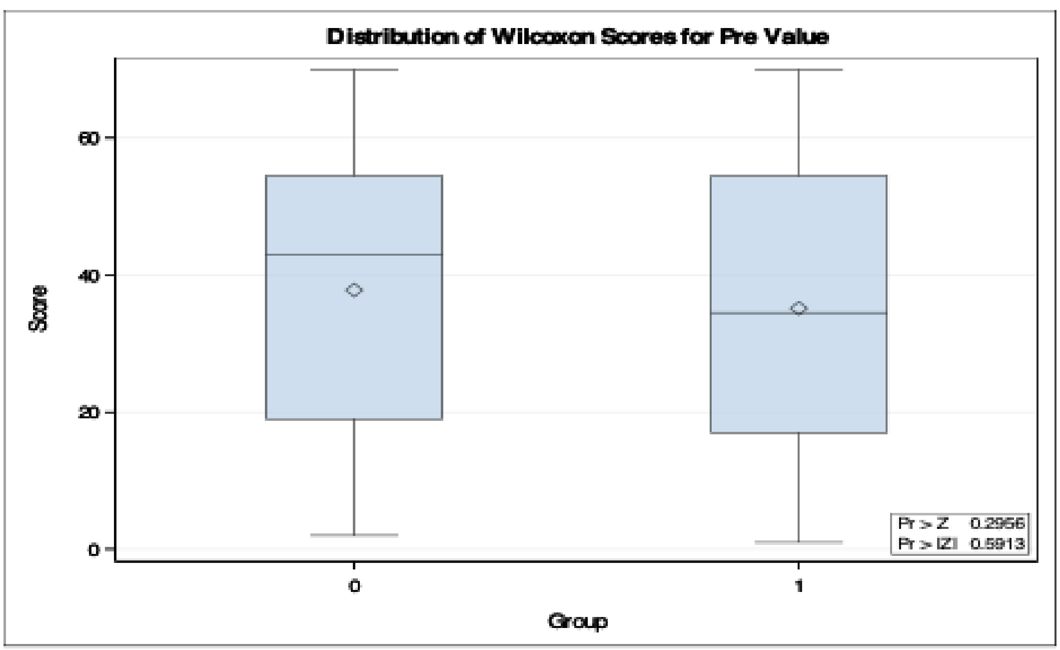




Wilcoxon Scores (Rank Sums) for Variable Pre Value Classified by Variable Group					
Group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
0	36	1362.0	1314.0	88.454333	37.833333
1	36	1266.0	1314.0	88.454333	35.166667
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	Z	Pr > Z	Pr > |Z|	Pr > Z	Pr > |Z|
1362.000	0.5370	0.2956	0.5913	0.2965	0.5929
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
0.2945	1	0.5874



ANCOVA results comparing Treatment and Control Groups

Class Level Information		
Class	**Levels**	**Values**
Group	2	0 1

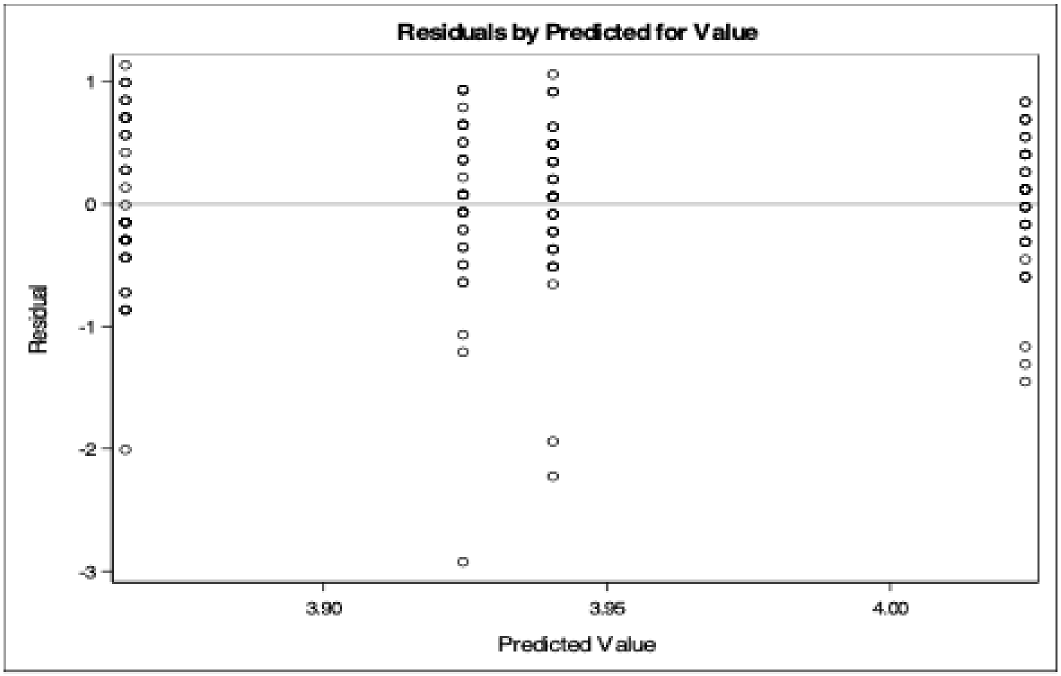
Number of Observations Read	297
Number of Observations Used	144

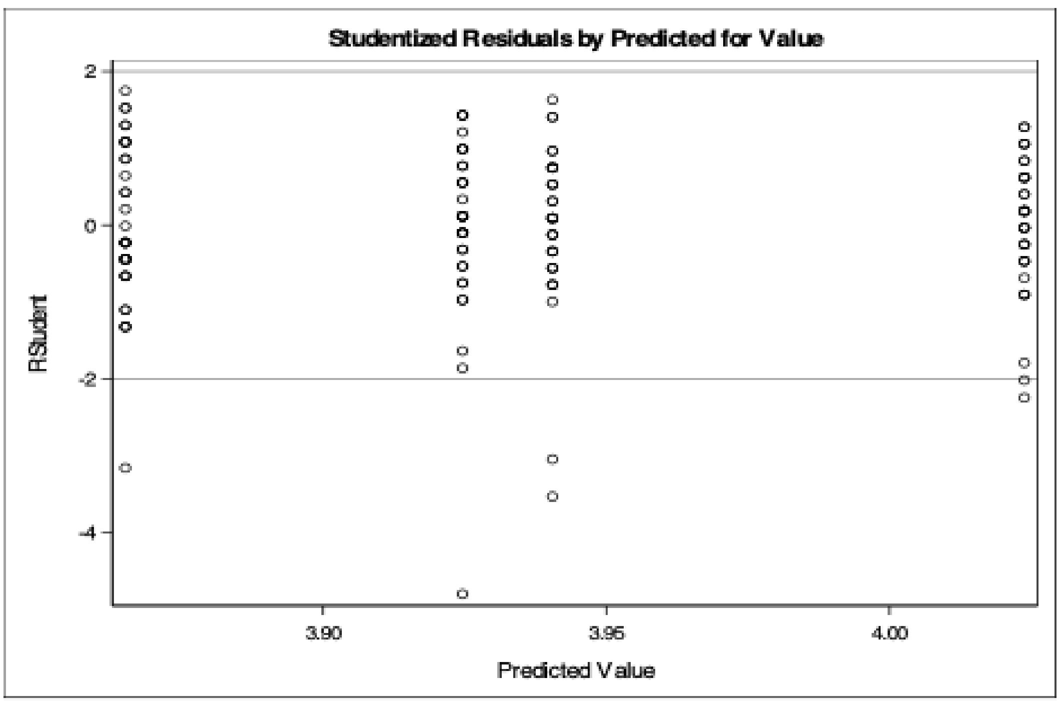
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.46315193	0.15438398	0.35	0.7894
Error	140	61.80839002	0.44148850		
Corrected Total	143	62.27154195			

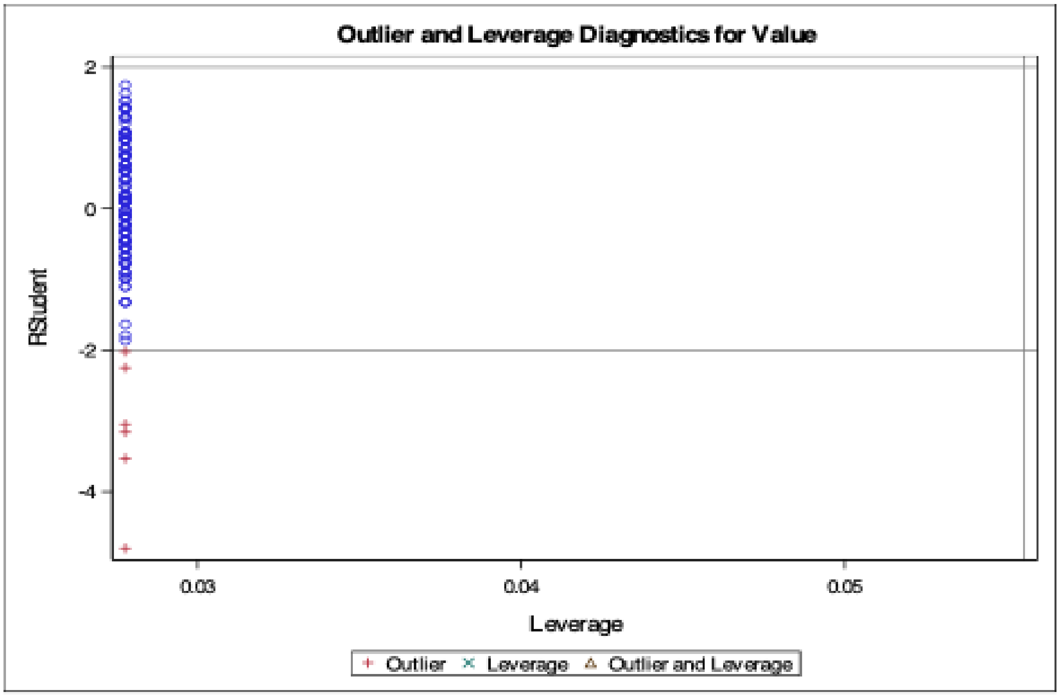
R-Square	Coeff Var	Root MSE	Value Mean
0.007438	16.87057	0.664446	3.938492

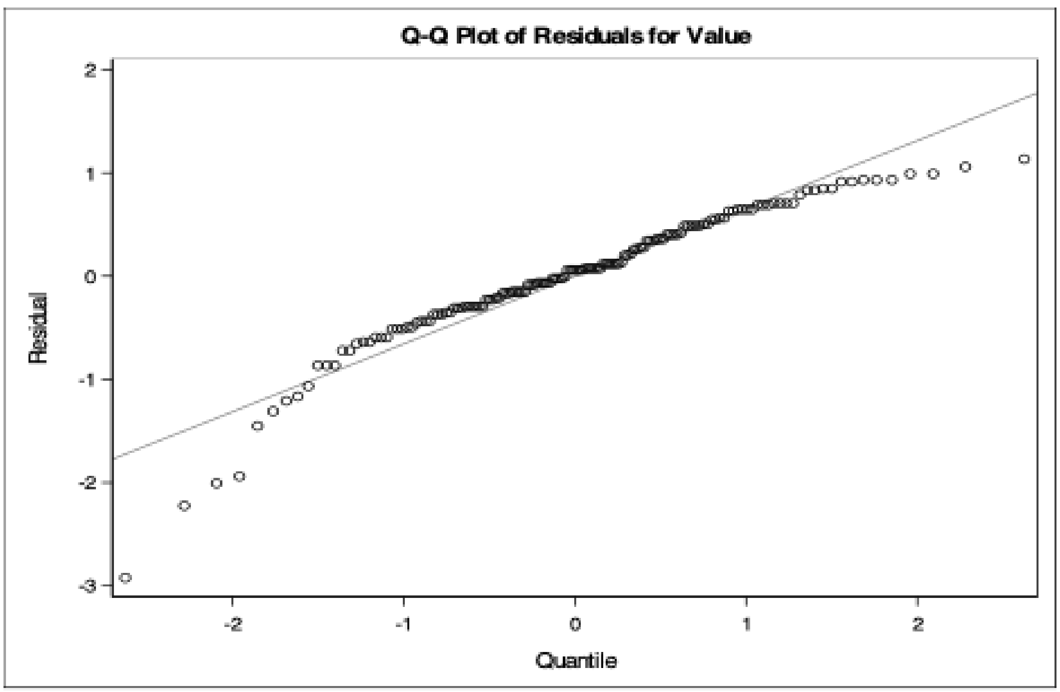
Source	DF	Type I SS	Mean Square	F Value	Pr > F
Group	1	0.00510204	0.00510204	0.01	0.9145
Pre-Post	1	0.18367347	0.18367347	0.42	0.5200
Pre-Post*Group	1	0.27437642	0.27437642	0.62	0.4318

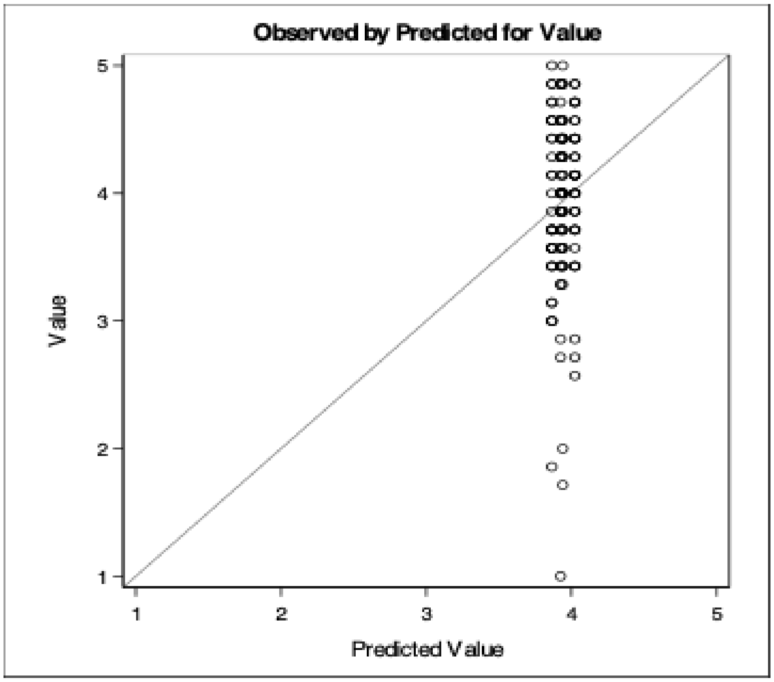
Source	DF	Type III SS	Mean Square	F Value	Pr > F
Group	1	0.17715420	0.17715420	0.40	0.5275
Pre-Post	1	0.18367347	0.18367347	0.42	0.5200
Pre-Post*Group	1	0.27437642	0.27437642	0.62	0.4318

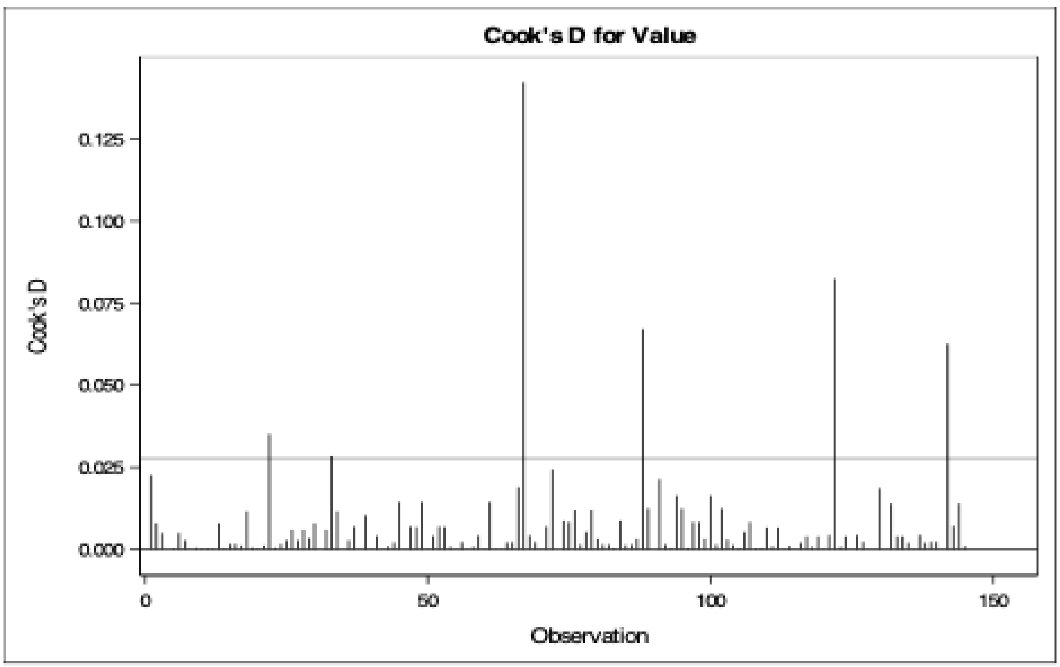

Residuals by Predicted for Value
Residual
1
0
-1
-2
-3
3.90
3.95
4.00
Predicted Value

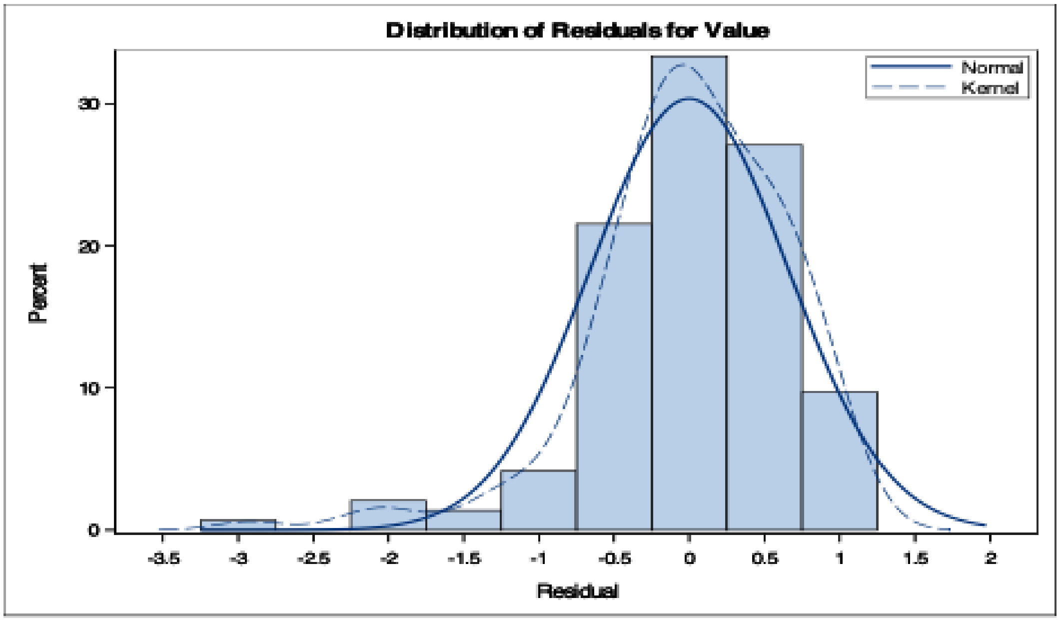

Studentized Residuals by Predicted for Value
RStudent
2
0
-2
-4
3.90
3.95
4.00
Predicted Value

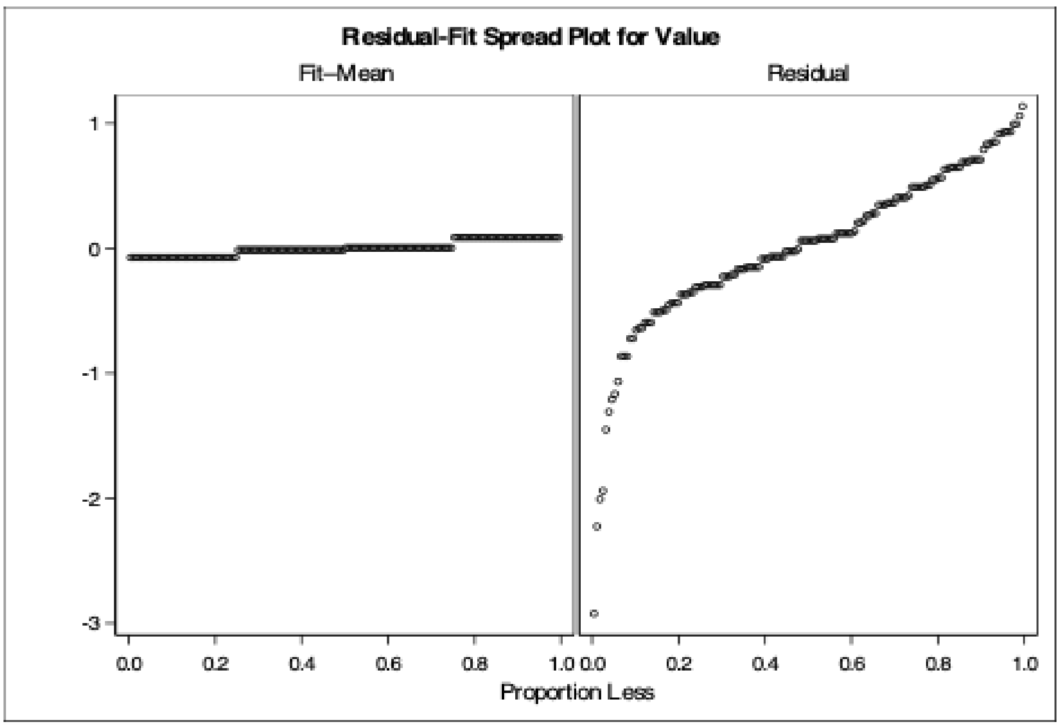

Outlier and Leverage Diagnostics for Value
RStudent
2
0
-2
-4
0.03
0.04
0.05
Leverage
Outlier
Leverage
Outlier and Leverage

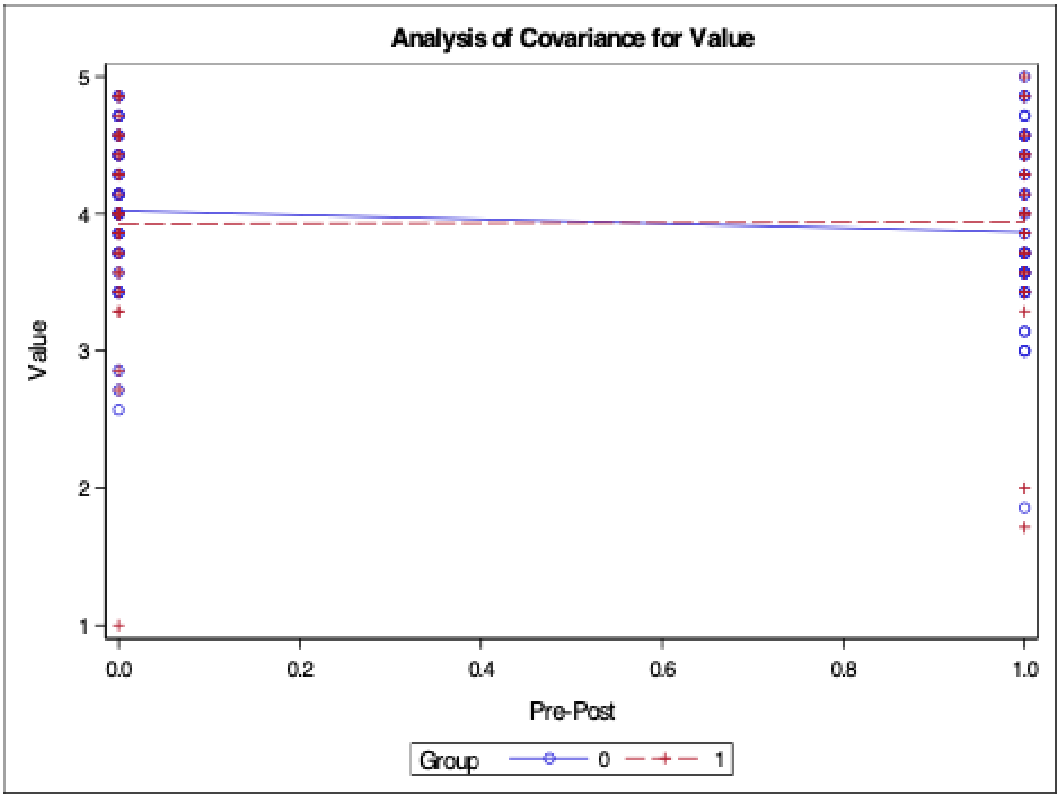

Q-Q Plot of Residuals for Value
Residual
2
1
0
-1
-2
-3
-2
-1
0
1
2
Quantile


Observed by Predicted for Value
Value
1
2
3
4
5
Predicted Value

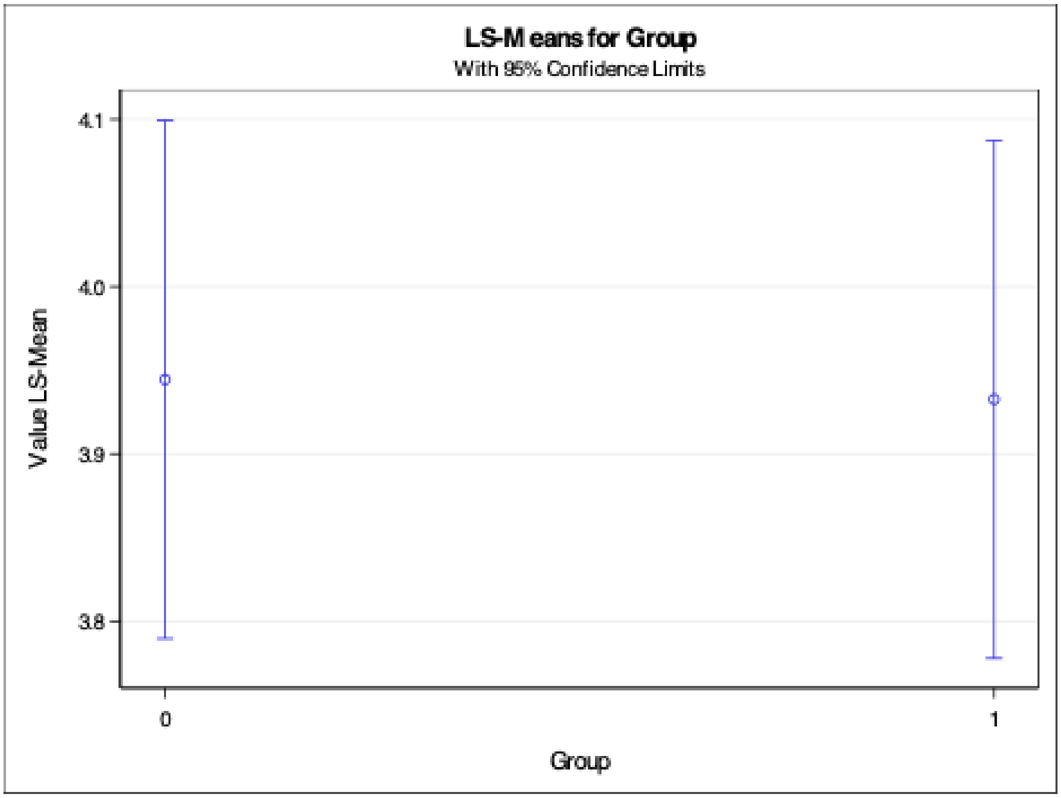

Cook's D for Value
Cook's D
0.000
0.025
0.050
0.075
0.100
0.125
0
50
100
150
Observation


Distribution of Residuals for Value
Normal
Kernel
Percent
Residual
30
20
10
0
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2


Residual-Fit Spread Plot for Value
Fit–Mean
Residual
Proportion Less
1
0
-1
-2
-3
0.0
0.2
0.4
0.6
0.8
1.0



Group	Value LSMEAN	H0:LSMean1=LSMean2
		Pr > |t|
0	3.94444444	0.9145
1	3.93253968	


LS-Means for Group
With 95% Confidence Limits
Value LS-Mean
4.1
4.0
3.9
3.8
0
1
Group

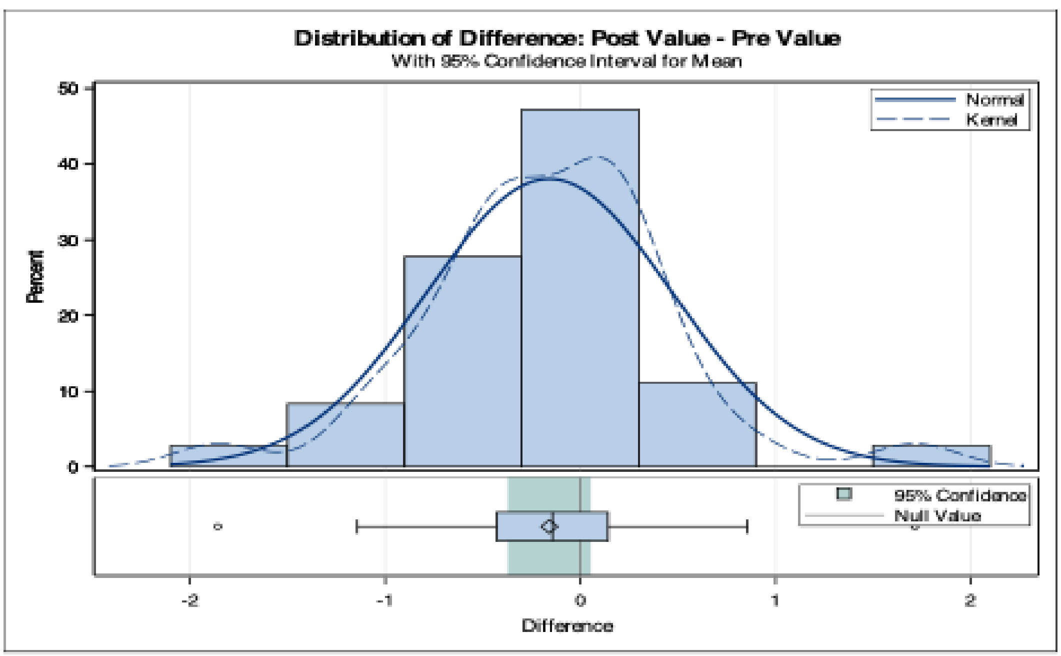
Pre-post Paired T-test Control Group

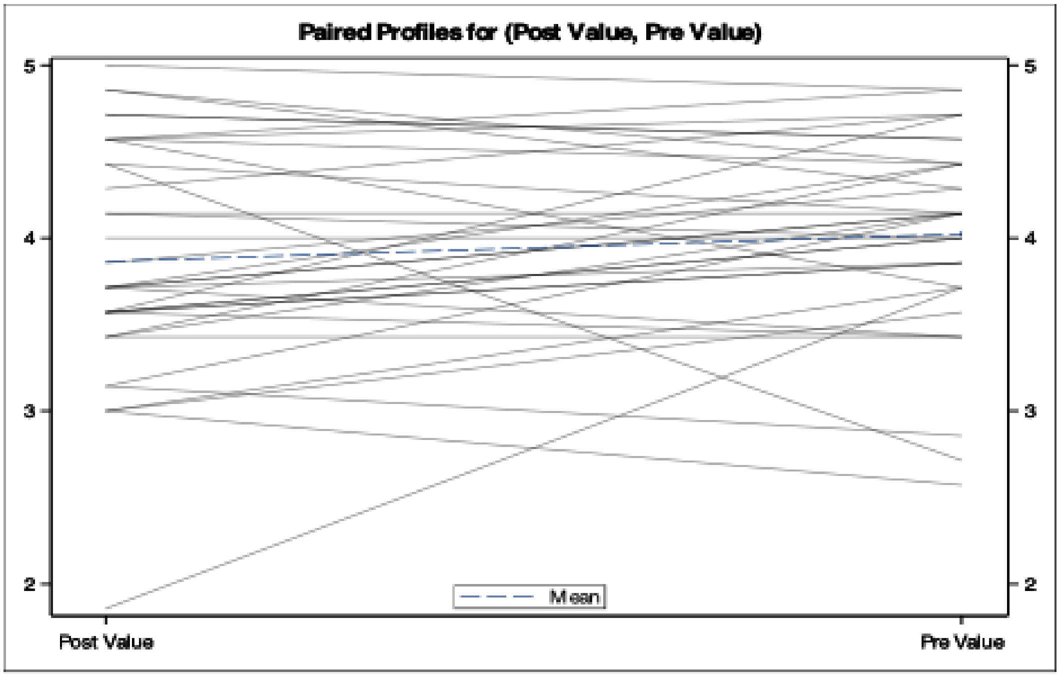
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.960663	**Pr < W**	0.2253
Kolmogorov-Smirnov	**D**	0.111628	**Pr > D**	>0.1500
Cramer-von Mises	**W-Sq**	0.073712	**Pr > W-Sq**	0.2457
Anderson-Darling	**A-Sq**	0.499041	**Pr > A-Sq**	0.2057

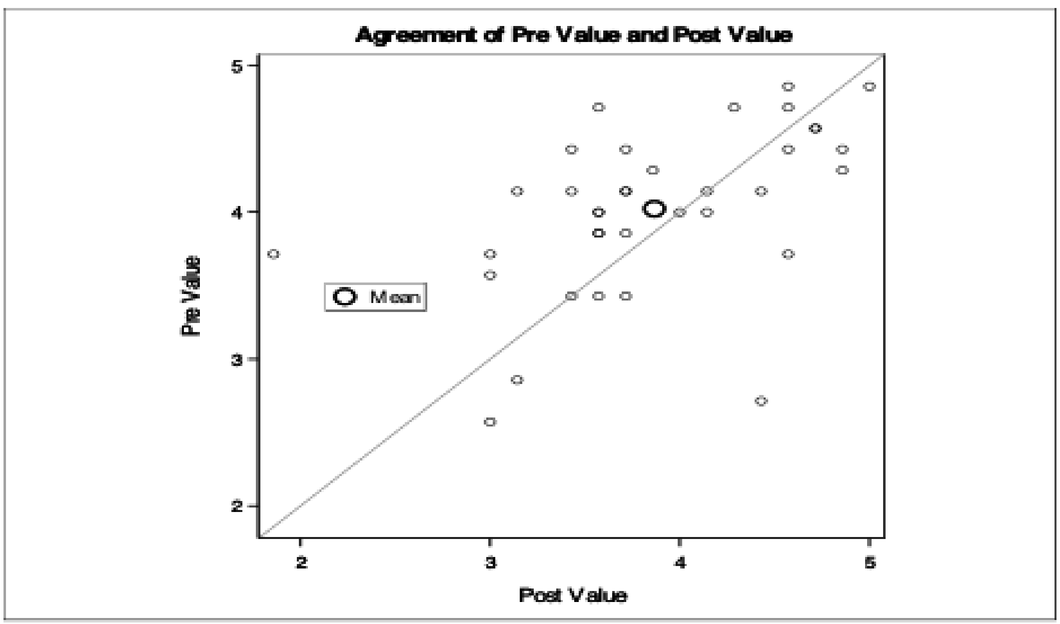
N	Mean	Std Dev	Std Err	Minimum	Maximum
36	-0.1587	0.6286	0.1048	-1.8571	1.7143

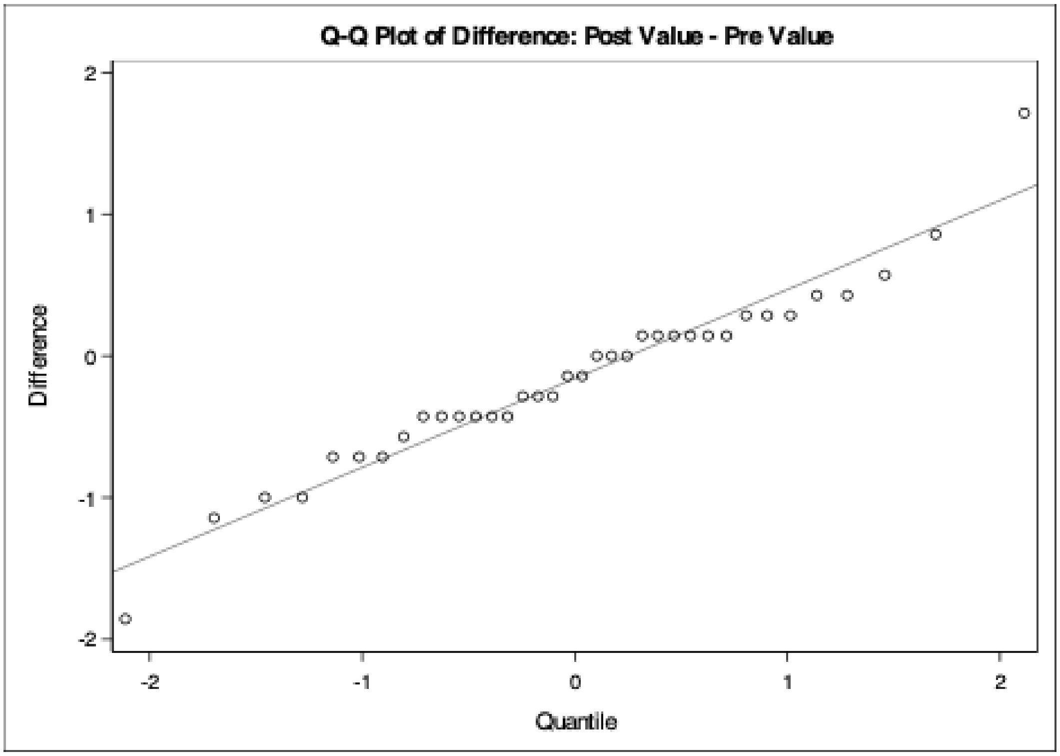
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-0.1587	-0.3714	0.0539	0.6286	0.5098	0.8199

DF	t Value	Pr > |t|
35	-1.52	0.1387




Paired Profiles for (Post Value, Pre Value)
5
4
3
2
Mean
Post Value
Pre Value


Agreement of Pre Value and Post Value
Mean
Pre Value
Post Value
2
3
4
5



Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	-1.5152	**Pr > |t|**	0.1387
Sign	**M**	-2.5	**Pr >= |M|**	0.4869
Signed Rank	**S**	-98.5	**Pr >= |S|**	0.0777

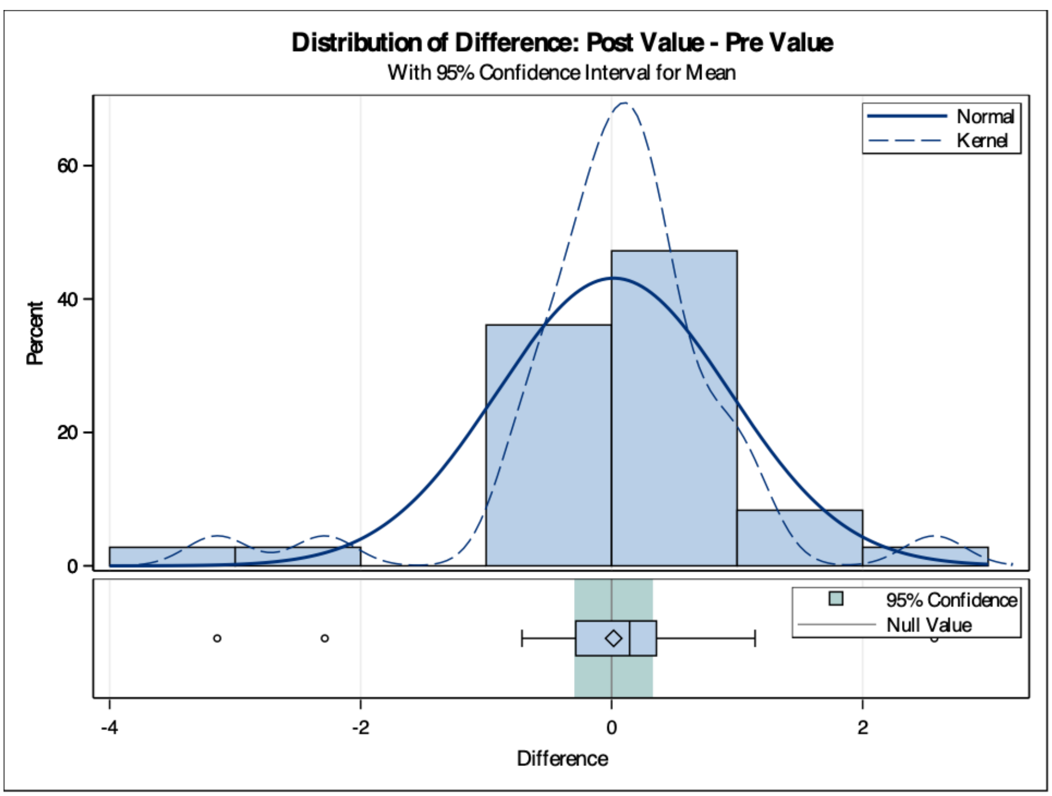
Pre-post Paired T-test Control Group

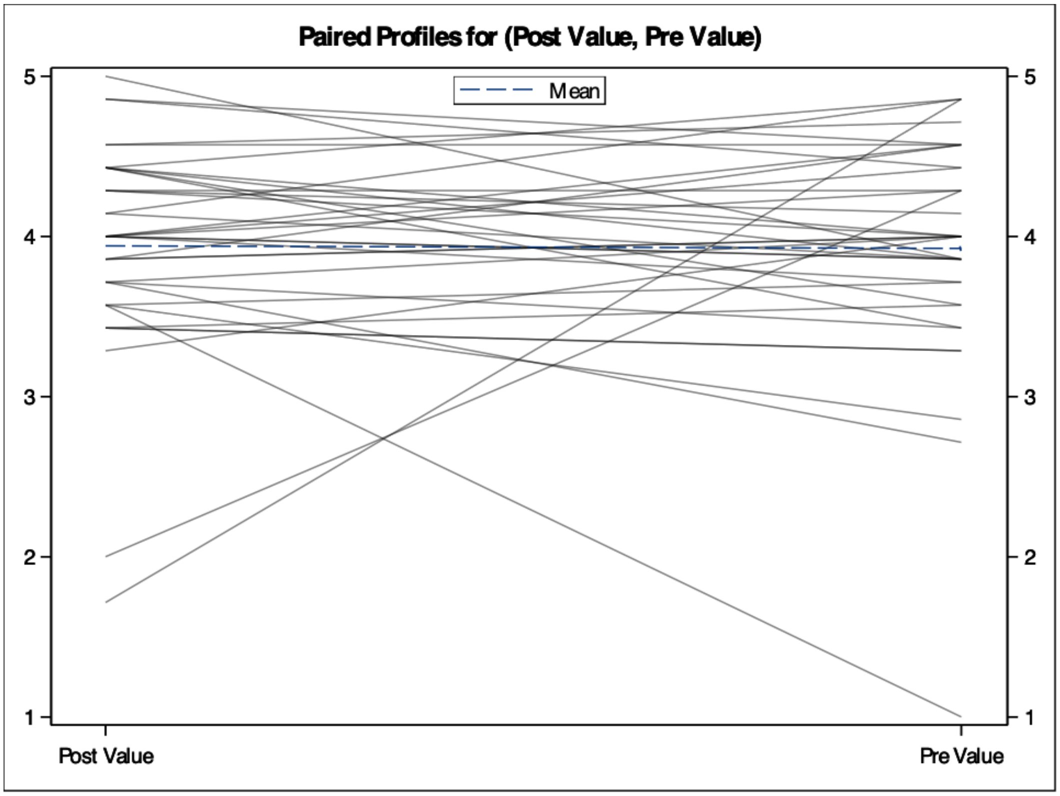
Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.866389	Pr < W	0.0005
Kolmogorov-Smirnov	D	0.159407	Pr > D	0.0208
Cramer-von Mises	W-Sq	0.258011	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	1.602454	Pr > A-Sq	<0.0050

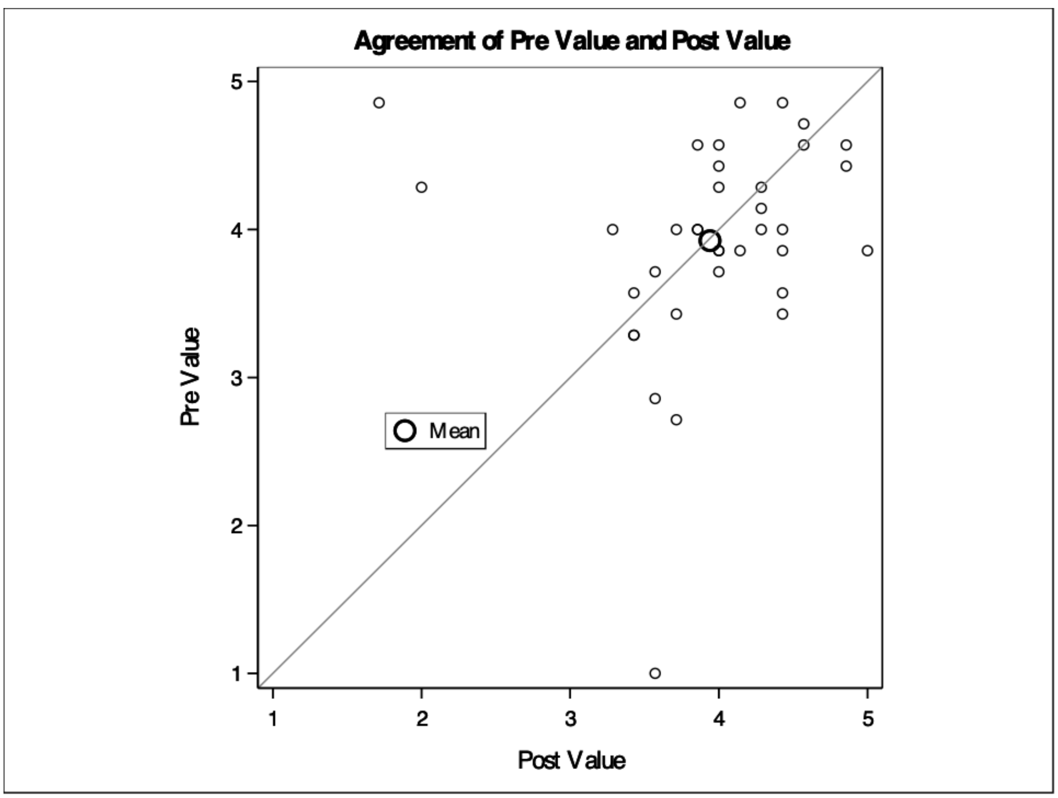
N	Mean	Std Dev	Std Err	Minimum	Maximum
36	0.0159	0.9251	0.1542	-3.1429	2.5714

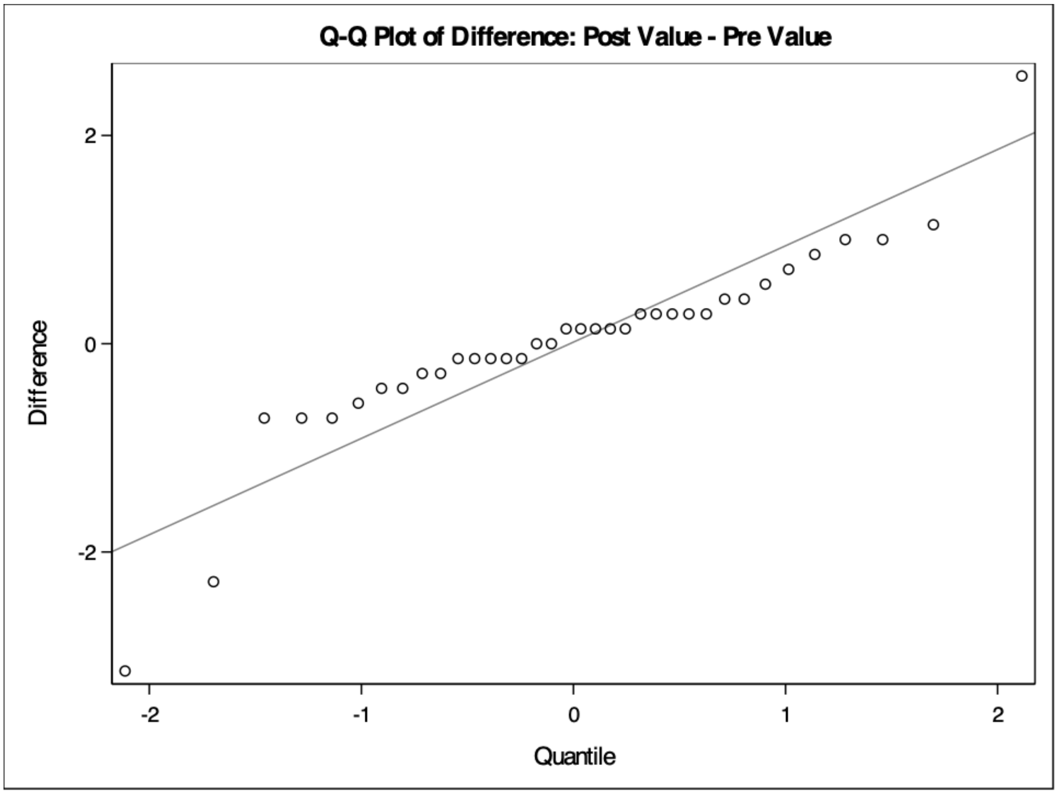
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0.0159	-0.2971	0.3289	0.9251	0.7503	1.2067

DF	t Value	Pr > |t|
35	0.10	0.9186




Paired Profiles for (Post Value, Pre Value)
Mean
5
4
3
2
1
Post Value
Pre Value


Agreement of Pre Value and Post Value
Mean
Pre Value
Post Value
1
2
3
4
5



Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	0.102955	Pr > |t|	0.9186
Sign	M	2	Pr >= |M|	0.6076
Signed Rank	S	34	Pr >= |S|	0.5684

Post-survey T-test comparing Treatment and Control Groups

Control Group (0)

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.945637	**Pr < W**	0.0762
Kolmogorov-Smirnov	**D**	0.143256	**Pr > D**	0.0609
Cramer-von Mises	**W-Sq**	0.10117	**Pr > W-Sq**	0.1055
Anderson-Darling	**A-Sq**	0.631008	**Pr > A-Sq**	0.0941

Treatment Group (1)

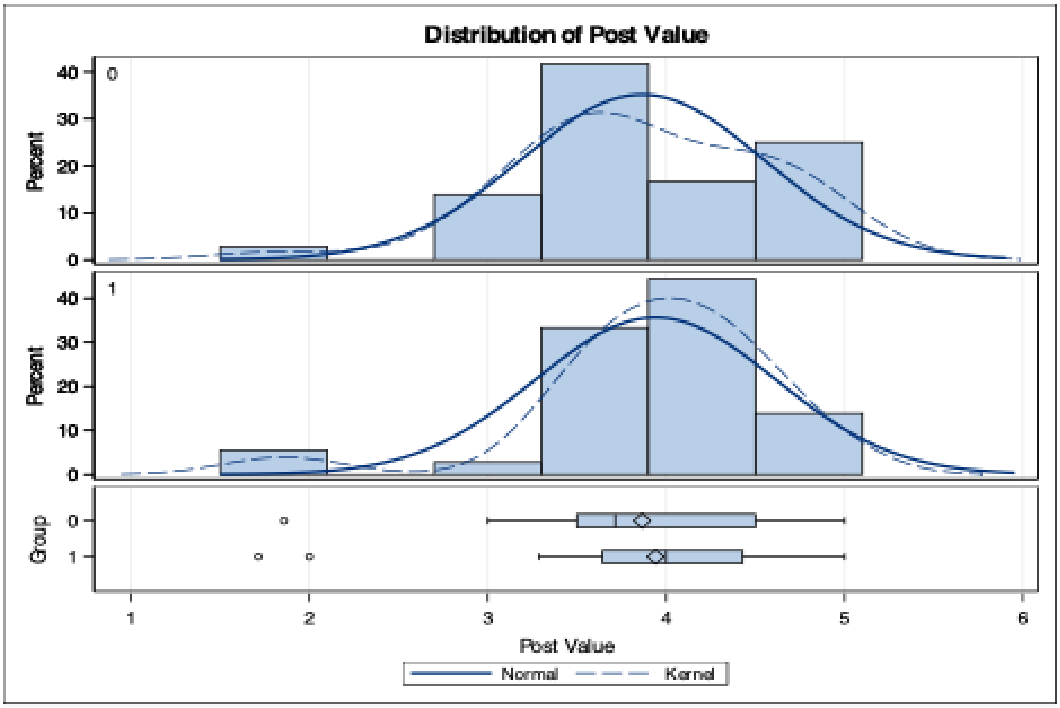
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.871957	**Pr < W**	0.0006
Kolmogorov-Smirnov	**D**	0.139515	**Pr > D**	0.0760
Cramer-von Mises	**W-Sq**	0.15033	**Pr > W-Sq**	0.0228
Anderson-Darling	**A-Sq**	1.147064	**Pr > A-Sq**	<0.0050

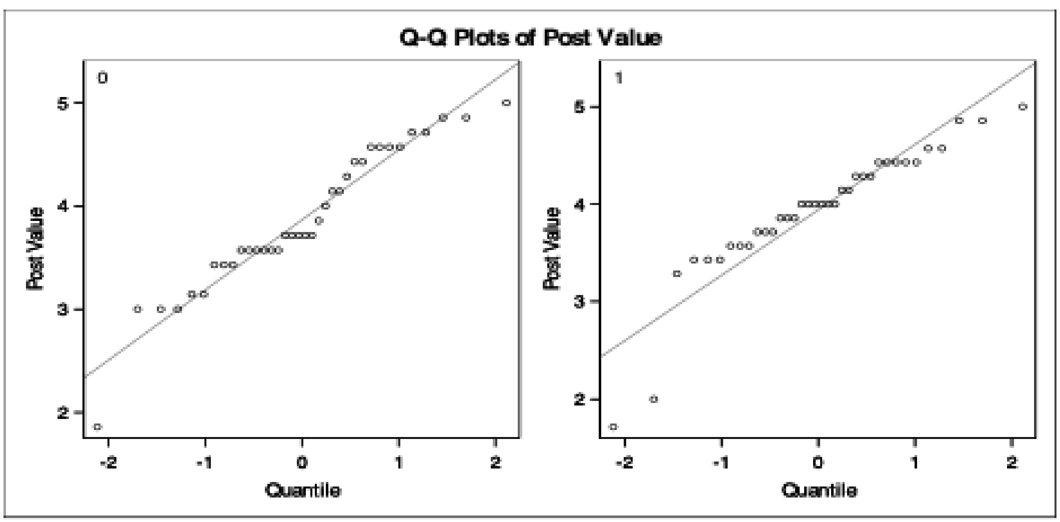
Group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		36	3.8651	0.6804	0.1134	1.8571	5.0000
1		36	3.9405	0.6713	0.1119	1.7143	5.0000
Diff (1-2)	**Pooled**		-0.0754	0.6758	0.1593		
Diff (1-2)	**Satterthwaite**		-0.0754		0.1593		

Group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3.8651	3.6349	4.0953	0.6804	0.5518	0.8875
1		3.9405	3.7134	4.1676	0.6713	0.5444	0.8756
Diff (1-2)	**Pooled**	-0.0754	-0.3931	0.2423	0.6758	0.5801	0.8098
Diff (1-2)	**Satterthwaite**	-0.0754	-0.3931	0.2423			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	70	-0.47	0.6375
Satterthwaite	Unequal	69.987	-0.47	0.6375

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	35	35	1.03	0.9368

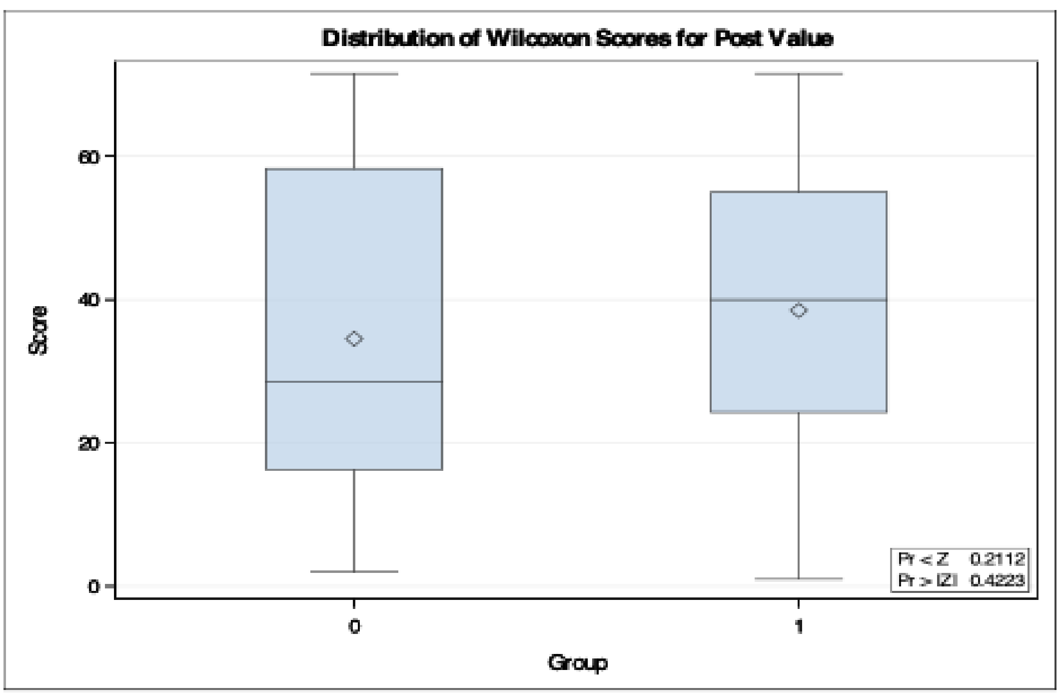




Wilcoxon Scores (Rank Sums) for Variable Post Value Classified by Variable Group					
Group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
0	36	1242.50	1314.0	88.482273	34.513889
1	36	1385.50	1314.0	88.482273	38.486111
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	Z	Pr < Z	Pr > |Z|	Pr < Z	Pr > |Z|
1242.500	-0.8024	0.2112	0.4223	0.2125	0.4250
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
0.6530	1	0.4190



Appendix H - Statistical Analysis for ATMI Subscale Enjoyment

Pre-survey T-test comparing Treatment and Control Groups

Control Group (0)

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.970726	Pr < W	0.4457
Kolmogorov-Smirnov	D	0.131569	Pr > D	0.1139
Cramer-von Mises	W-Sq	0.108219	Pr > W-Sq	0.0870
Anderson-Darling	A-Sq	0.555236	Pr > A-Sq	0.1452

Treatment Group (1)

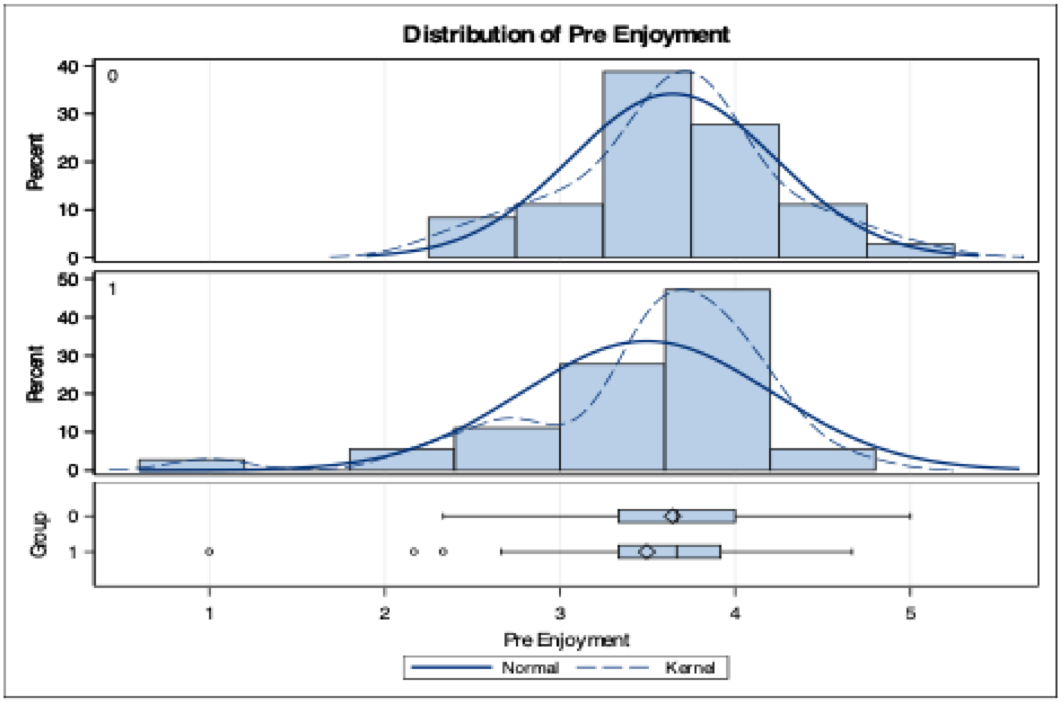
Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.882574	Pr < W	0.0012
Kolmogorov-Smirnov	D	0.224828	Pr > D	<0.0100
Cramer-von Mises	W-Sq	0.259983	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	1.376031	Pr > A-Sq	<0.0050

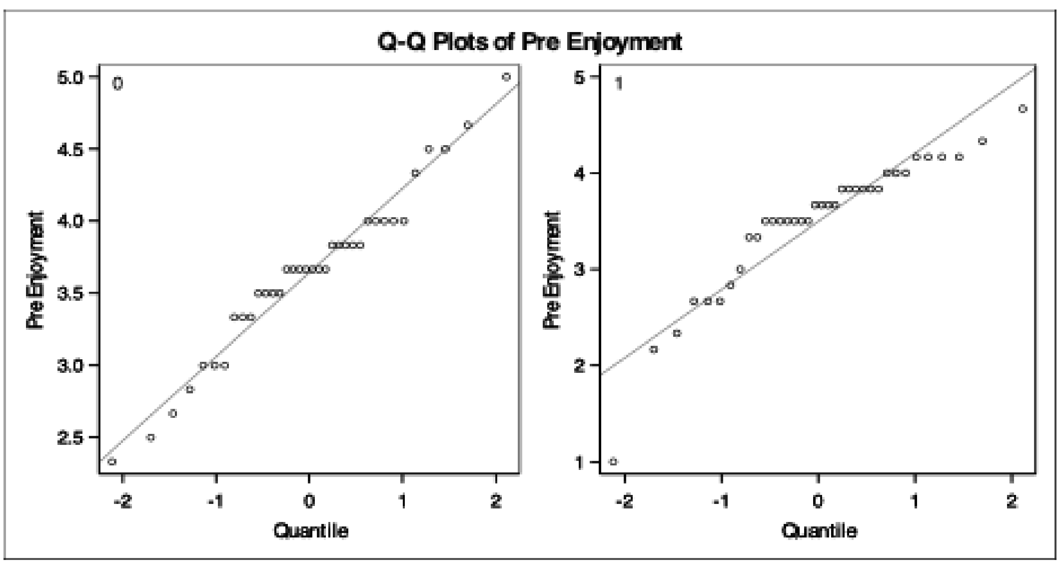
Group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		36	3.6435	0.5830	0.0972	2.3333	5.0000
1		36	3.4954	0.7088	0.1181	1.0000	4.6667
Diff (1-2)	Pooled		0.1481	0.6490	0.1530		
Diff (1-2)	Satterthwaite		0.1481		0.1530		

Group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3.6435	3.4462	3.8408	0.5830	0.4729	0.7605
1		3.4954	3.2556	3.7352	0.7088	0.5749	0.9246
Diff (1-2)	Pooled	0.1481	-0.1569	0.4532	0.6490	0.5570	0.7776
Diff (1-2)	Satterthwaite	0.1481	-0.1571	0.4534			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	70	0.97	0.3361
Satterthwaite	Unequal	67.49	0.97	0.3362

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	35	35	1.48	0.2528

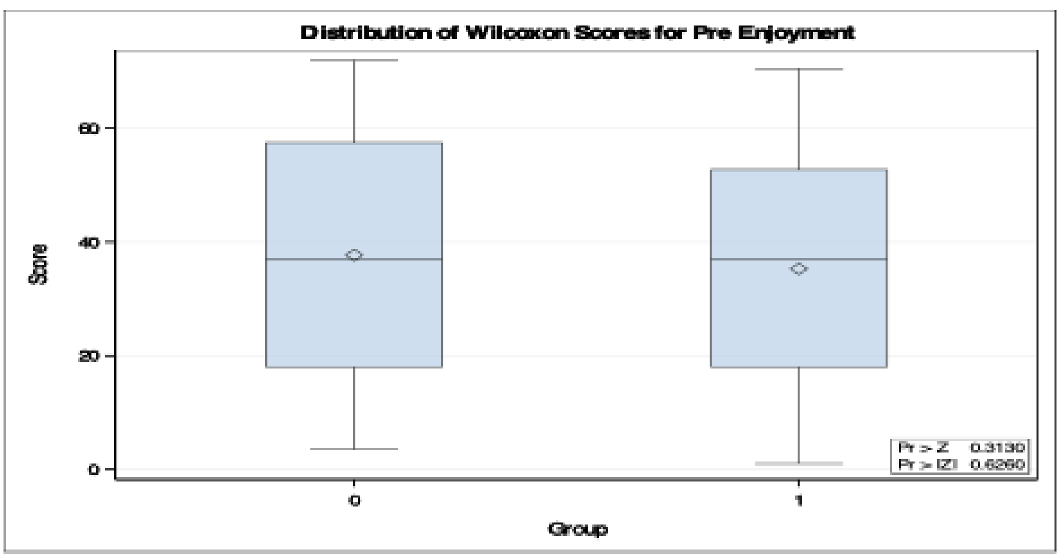




Wilcoxon Scores (Rank Sums) for Variable Pre Enjoyment Classified by Variable Group					
Group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
0	36	1357.50	1314.0	88.219716	37.708333
1	36	1270.50	1314.0	88.219716	35.291667
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	Z	Pr > Z	Pr > |Z|	Pr > Z	Pr > |Z|
1357.500	0.4874	0.3130	0.6260	0.3137	0.6275
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
0.2431	1	0.6220



ANCOVA results comparing Treatment and Control Groups

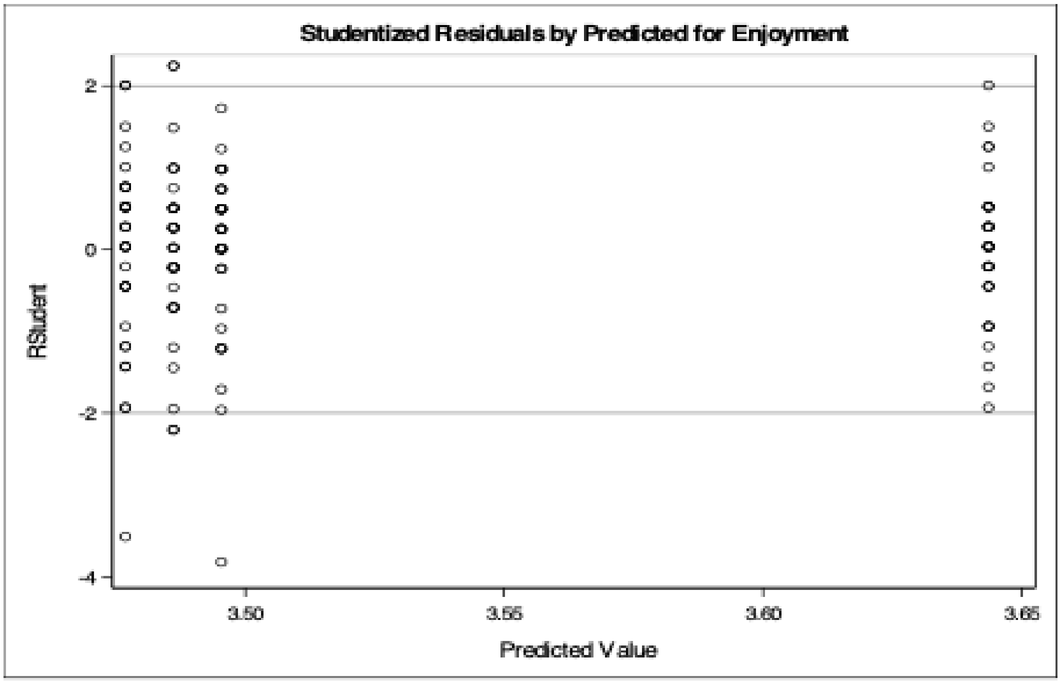
Class Level Information		
Class	**Levels**	**Values**
Group	2	0 1

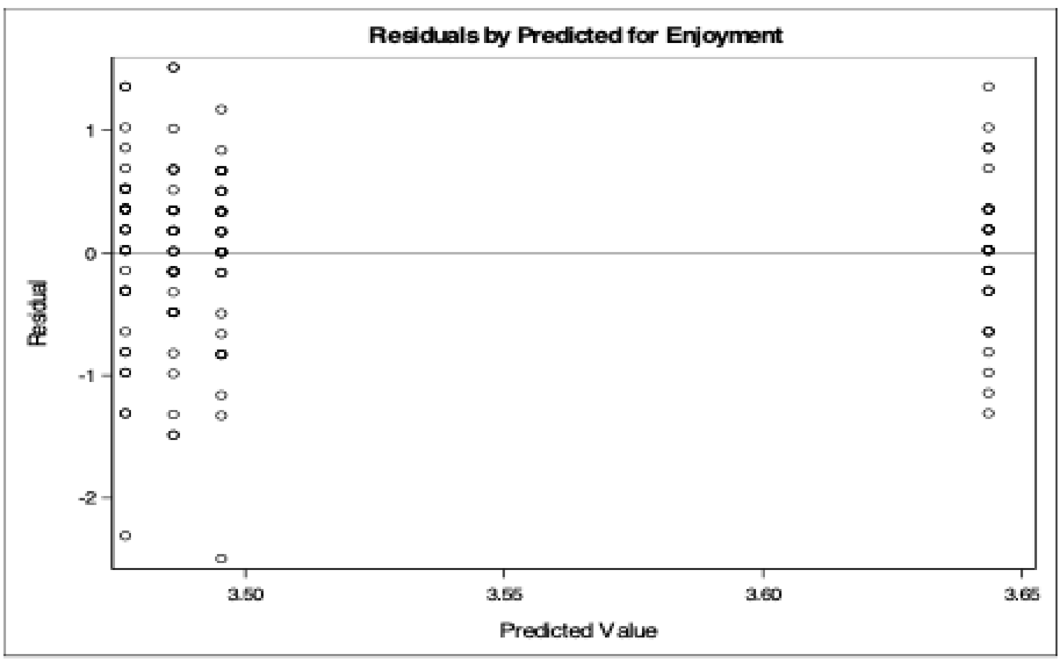
Number of Observations Read	297
Number of Observations Used	144

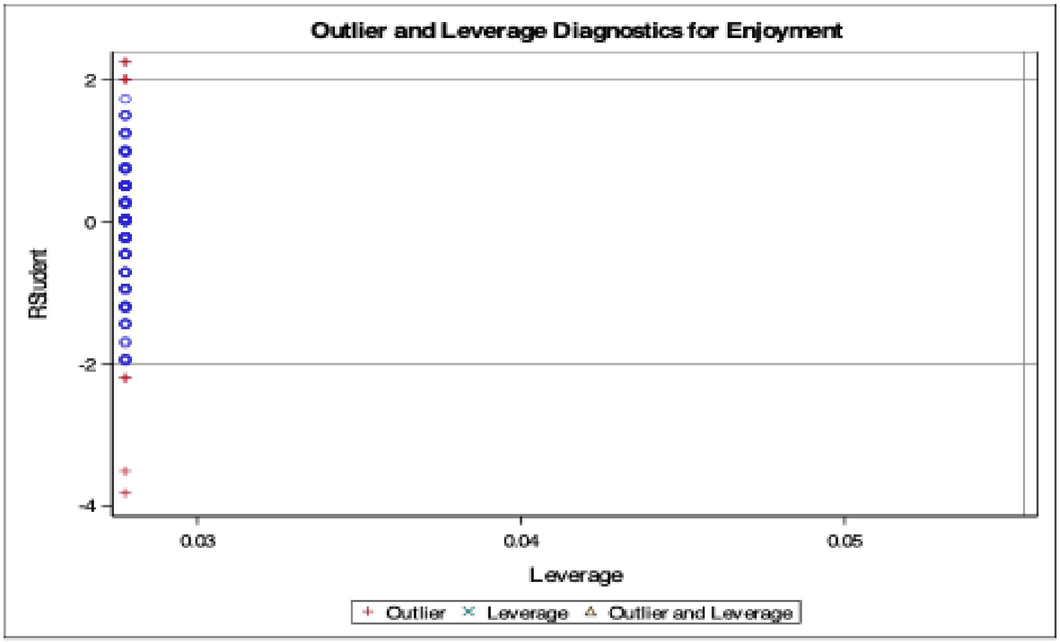
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	0.67515432	0.22505144	0.47	0.7052
Error	140	67.34259259	0.48101852		
Corrected Total	143	68.01774691			

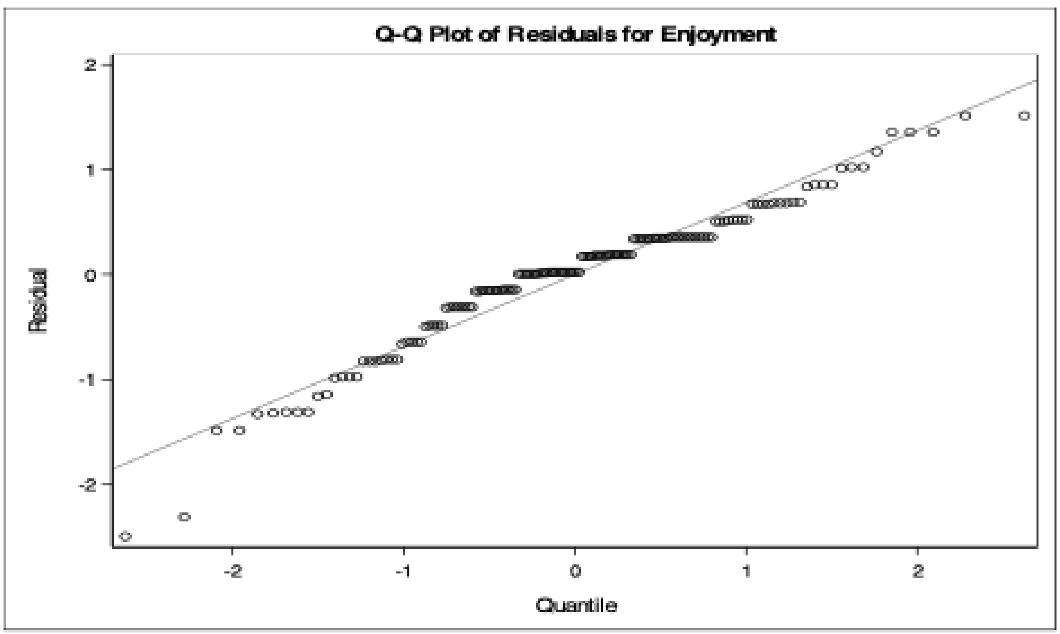
R-Square	Coeff Var	Root MSE	Enjoyment Mean
0.009926	19.67273	0.693555	3.525463

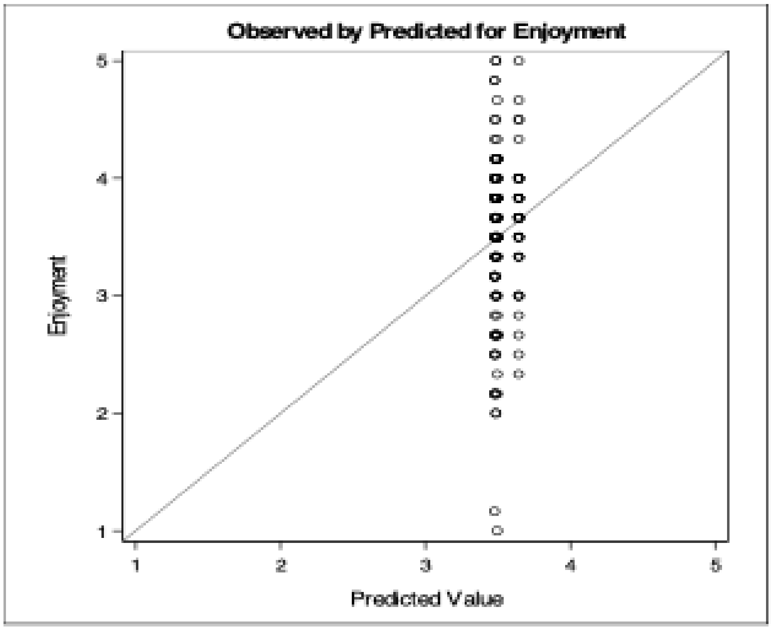
Source	DF	Type I SS	Mean Square	F Value	Pr > F
Group	1	0.22299383	0.22299383	0.46	0.4971
Pre-Post	1	0.27854938	0.27854938	0.58	0.4480
Pre-Post*Group	1	0.17361111	0.17361111	0.36	0.5490

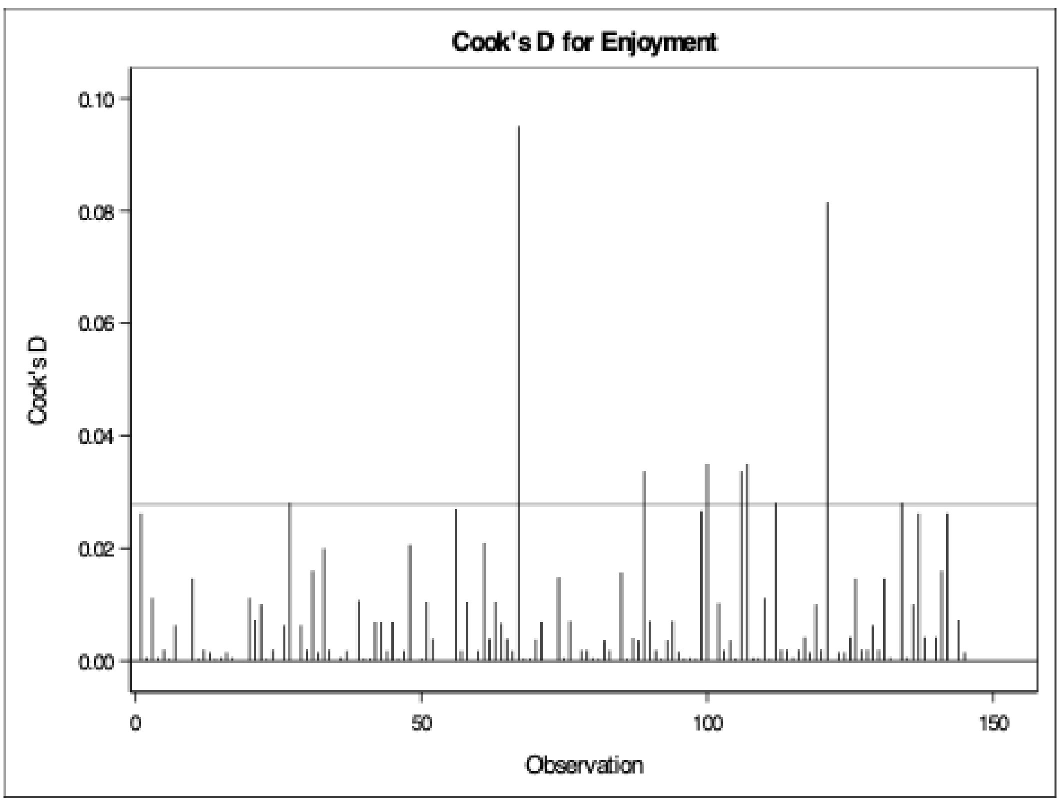

Studentized Residuals by Predicted for Enjoyment
RStudent
2
0
-2
-4
3.50
3.55
3.60
3.65
Predicted Value

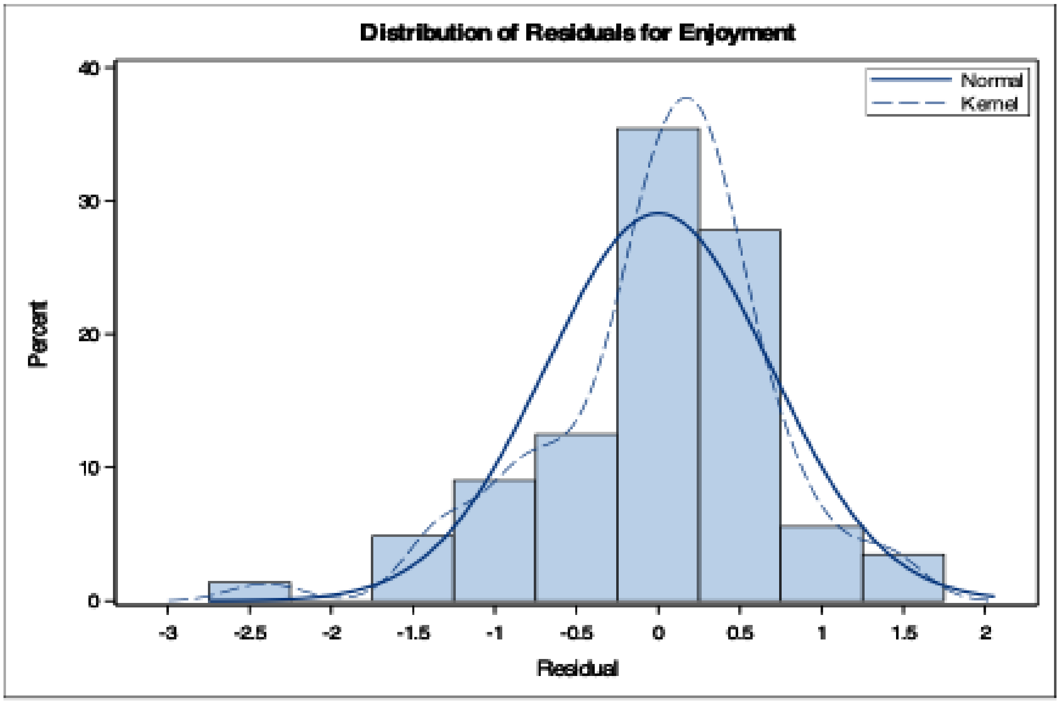

Residuals by Predicted for Enjoyment
Residual
1
0
-1
-2
3.50
3.55
3.60
3.65
Predicted Value

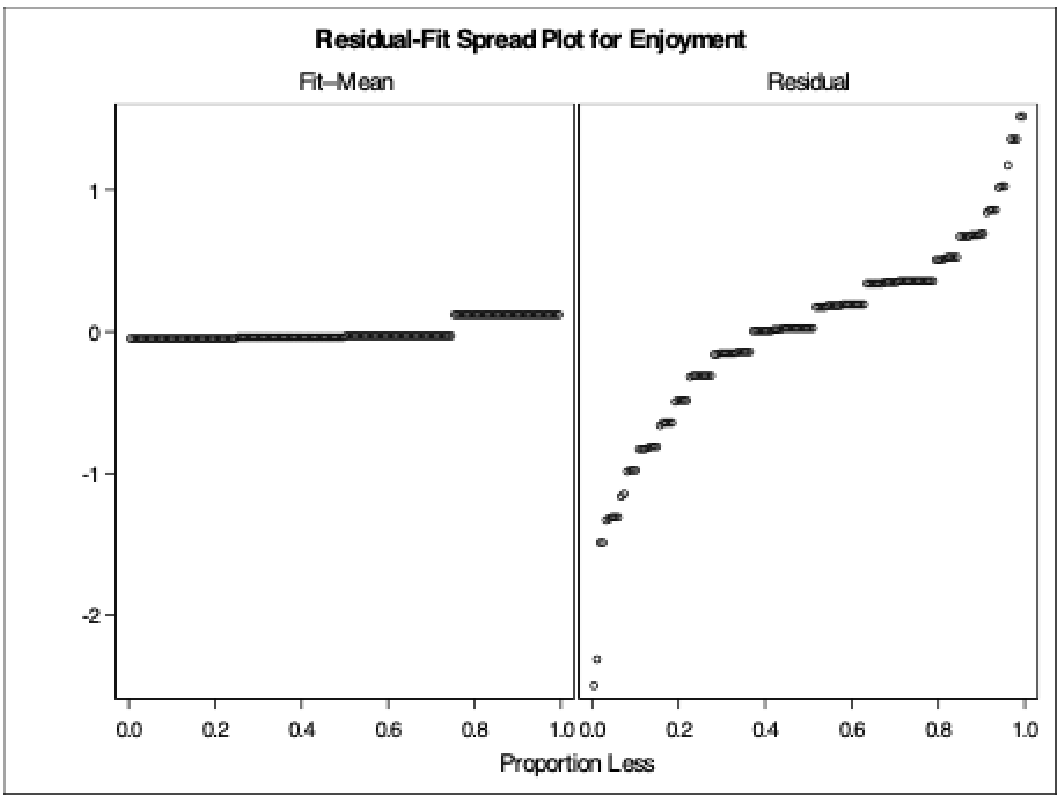

Outlier and Leverage Diagnostics for Enjoyment
RStudent
2
0
-2
-4
0.03
0.04
0.05
Leverage
Outlier
Leverage
Outlier and Leverage

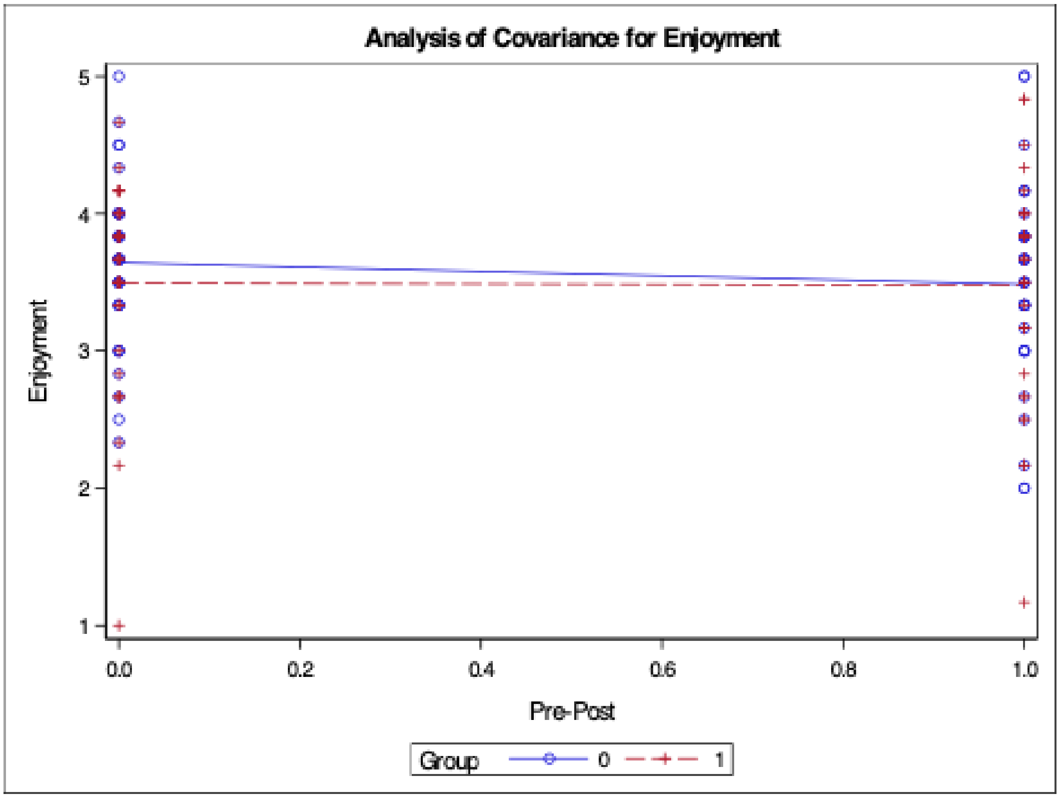

Q-Q Plot of Residuals for Enjoyment
Residual
2
1
0
-1
-2
-2
-1
0
1
2
Quantile


Observed by Predicted for Enjoyment
Enjoyment
Predicted Value
1
2
3
4
5

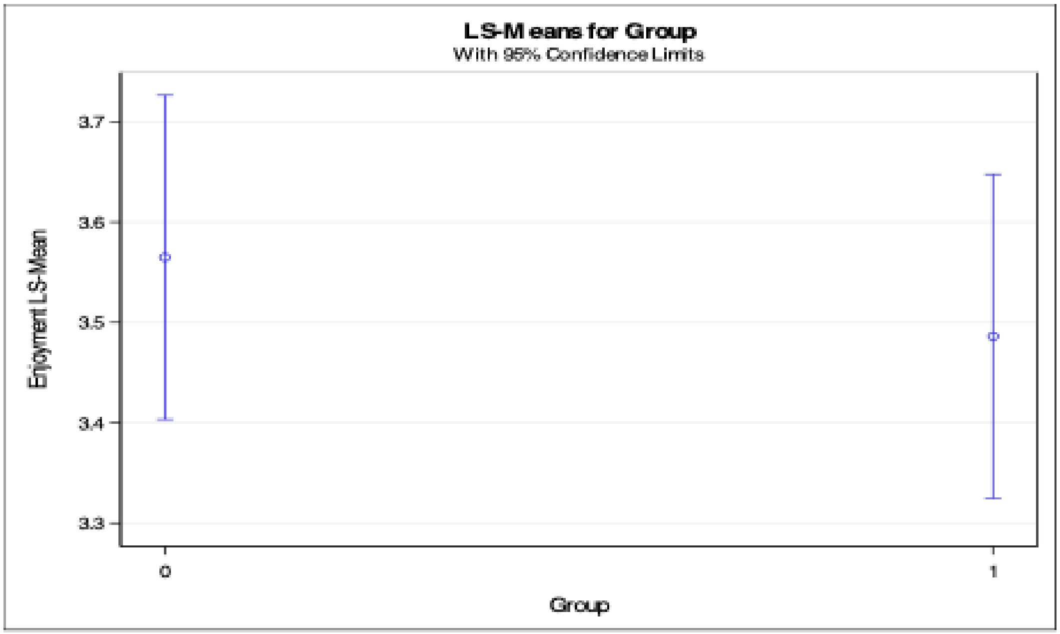

Cook's D for Enjoyment
Cook's D
0.10
0.08
0.06
0.04
0.02
0.00
Observation
0
50
100
150

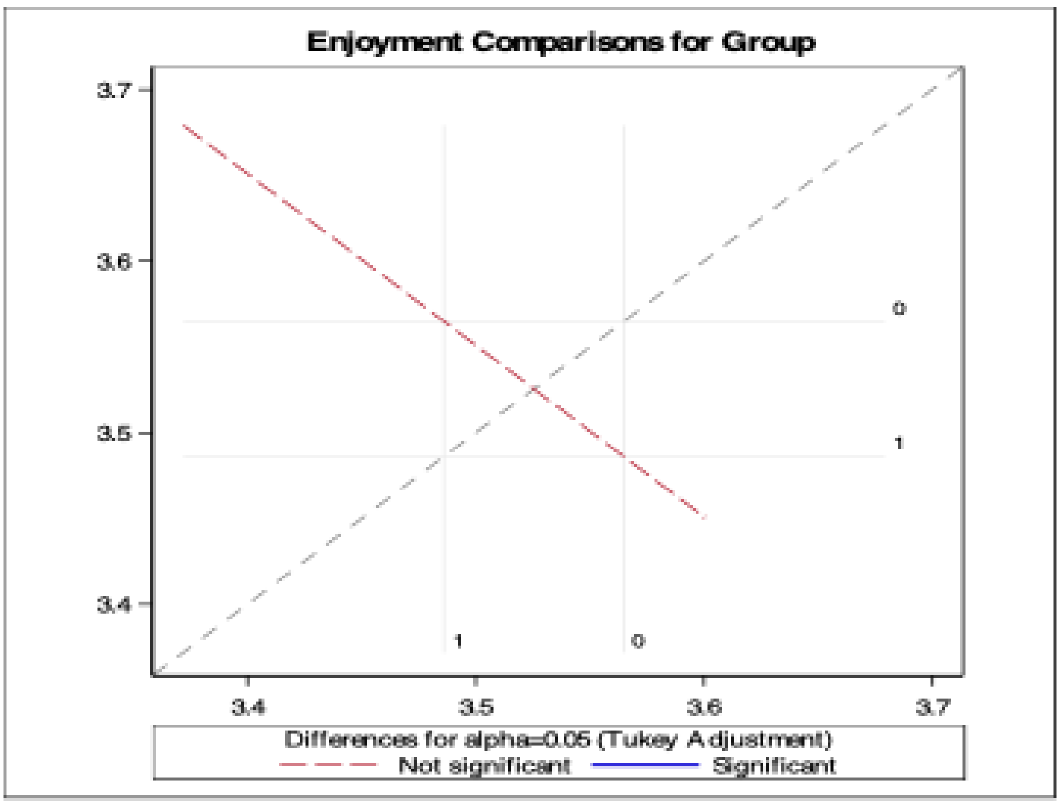

Distribution of Residuals for Enjoyment
Normal
Kernel
Percent
0
10
20
30
40
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
Residual


Residual-Fit Spread Plot for Enjoyment
Fit–Mean
Residual
1
0
-1
-2
0.0
0.2
0.4
0.6
0.8
1.0
Proportion Less



Group	Enjoyment LSMEAN	H0:LSMean1=LSMean2 Pr > |t|
0	3.56481481	0.4971
1	3.48611111	


LS-Means for Group
With 95% Confidence Limits
Enjoyment LS-Mean
3.7
3.6
3.5
3.4
3.3
0
1
Group


Enjoyment Comparisons for Group
3.7
3.6
3.5
3.4
0
1
1
0
3.4
3.5
3.6
3.7
Differences for alpha=0.05 (Tukey Adjustment)
Not significant
Significant

Pre-post Paired T-test Control Group

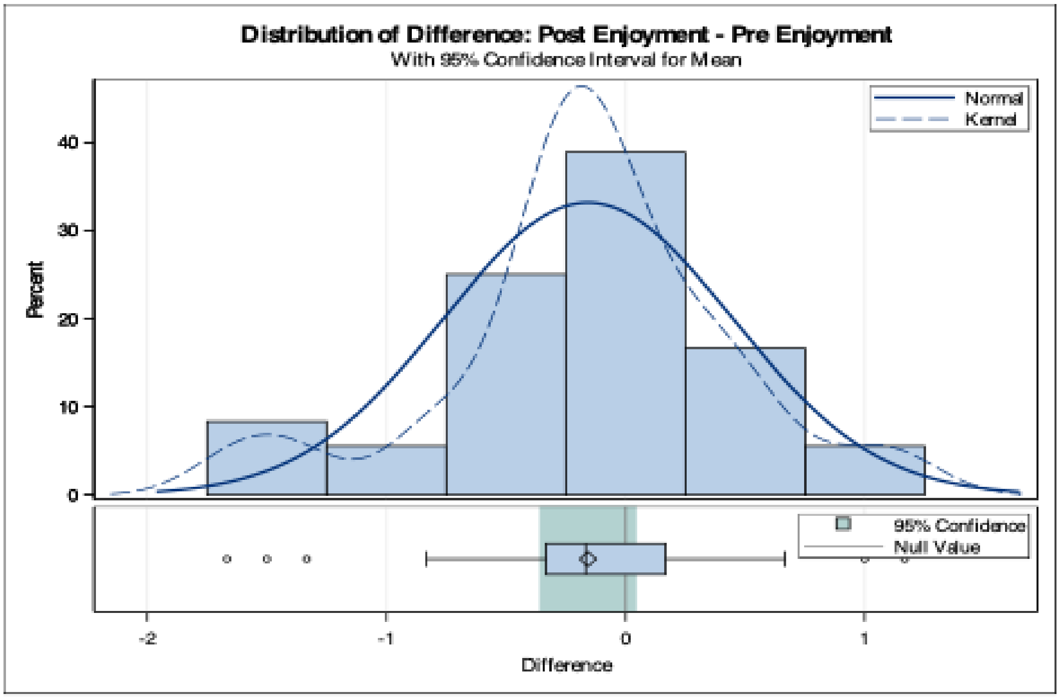
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.948812	**Pr < W**	0.0959
Kolmogorov-Smirnov	**D**	0.190518	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	0.154099	**Pr > W-Sq**	0.0209
Anderson-Darling	**A-Sq**	0.833218	**Pr > A-Sq**	0.0293

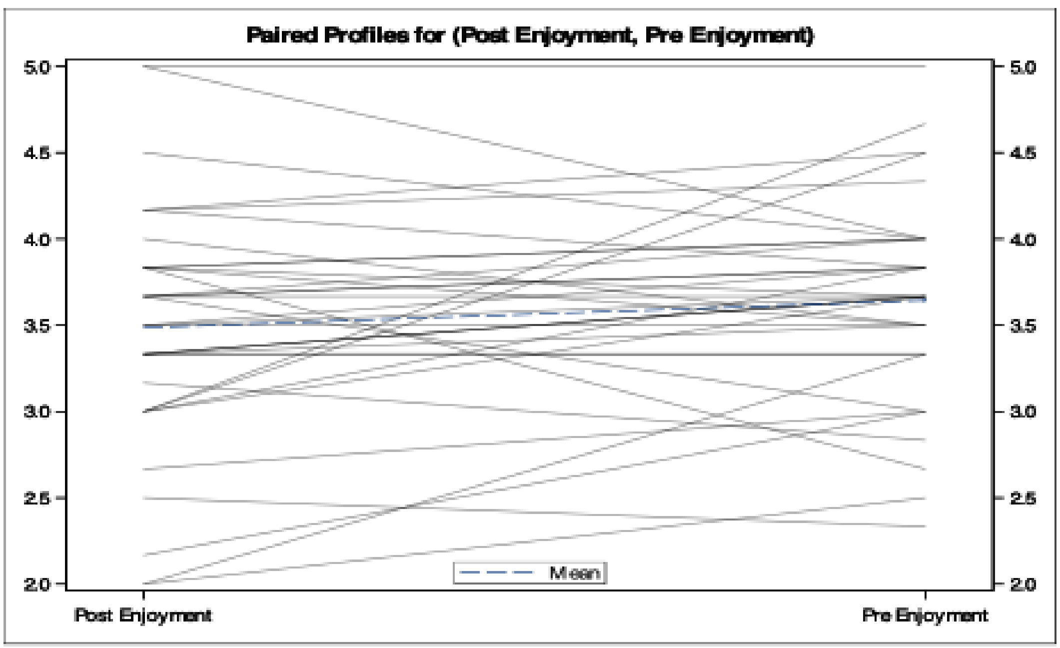
N	Mean	Std Dev	Std Err	Minimum	Maximum
36	-0.1574	0.6015	0.1003	-1.6667	1.1667

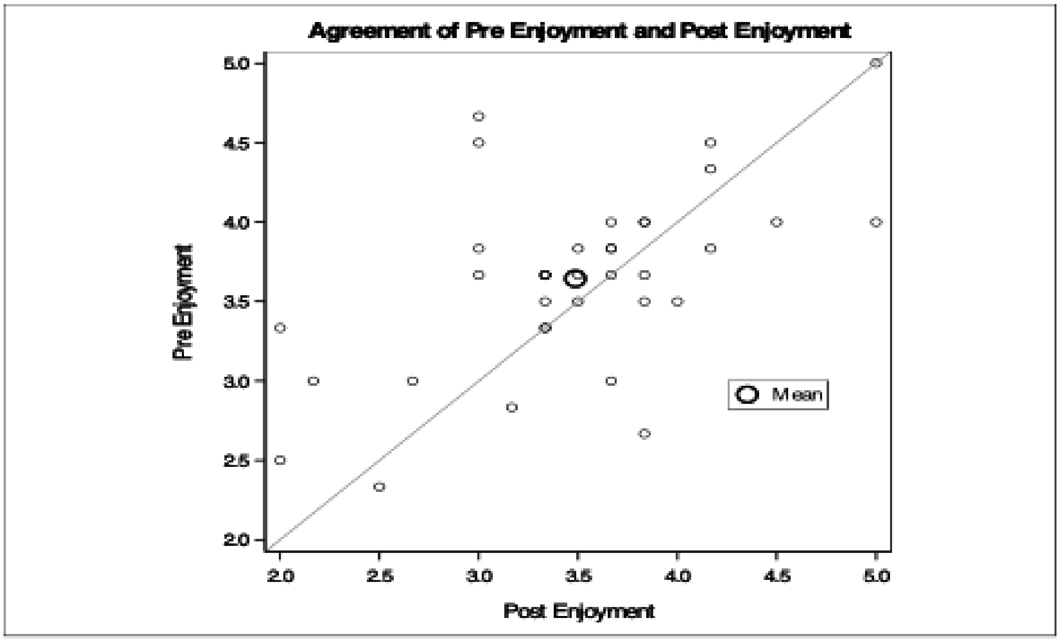
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-0.1574	-0.3609	0.0461	0.6015	0.4879	0.7846

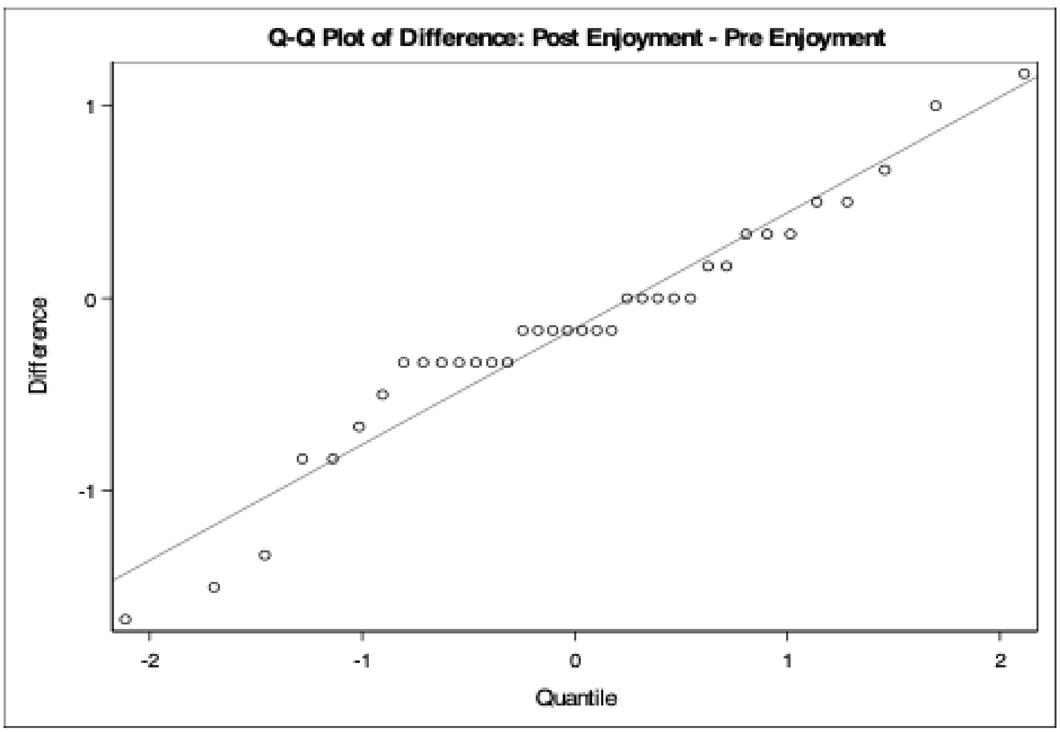
DF	t Value	Pr > |t|
35	-1.57	0.1254

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	-1.57012	**Pr > |t|**	0.1254
Sign	**M**	-5.5	**Pr >= |M|**	0.0708
Signed Rank	**S**	-75	**Pr >= |S|**	0.1436


Distribution of Difference: Post Enjoyment - Pre Enjoyment
With 95% Confidence Interval for Mean
Normal
Kernel
Percent
95% Confidence
Null Value
Difference


Paired Profiles for (Post Enjoyment, Pre Enjoyment)
Mean
Post Enjoyment
Pre Enjoyment





Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	-0.1513	Pr > |t|	0.8806
Sign	M	-0.5	Pr >= |M|	1.0000
Signed Rank	S	-18	Pr >= |S|	0.7037

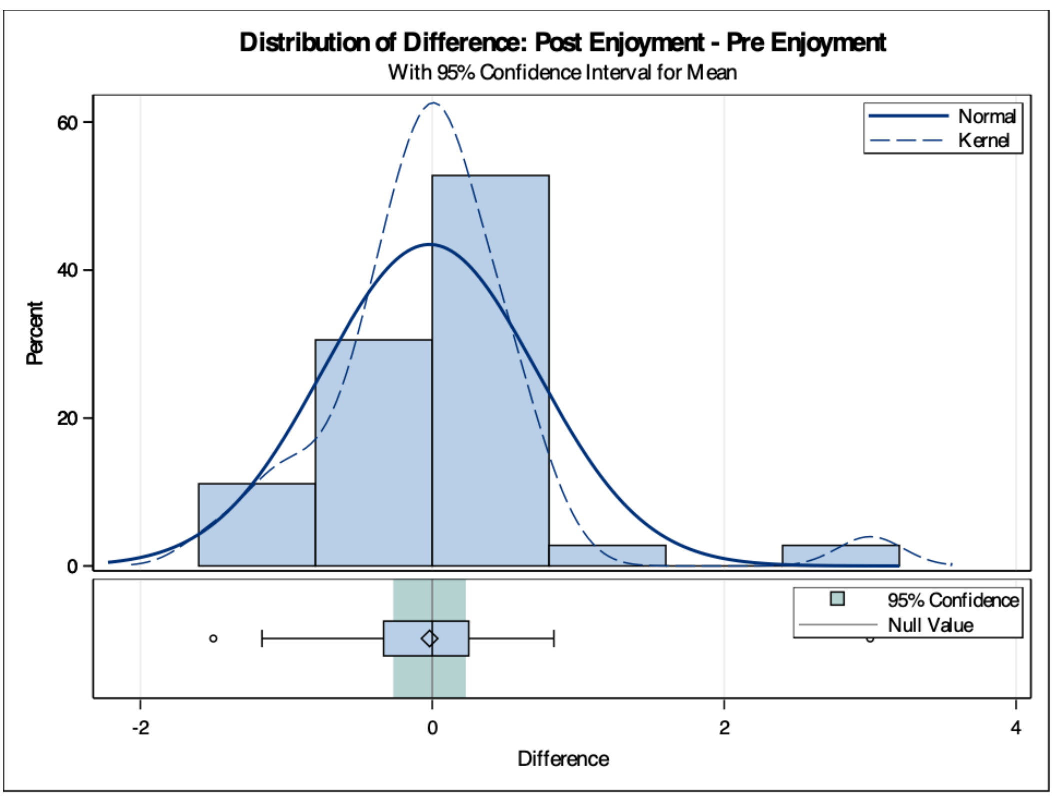
Pre-post Paired T-test Treatment Group

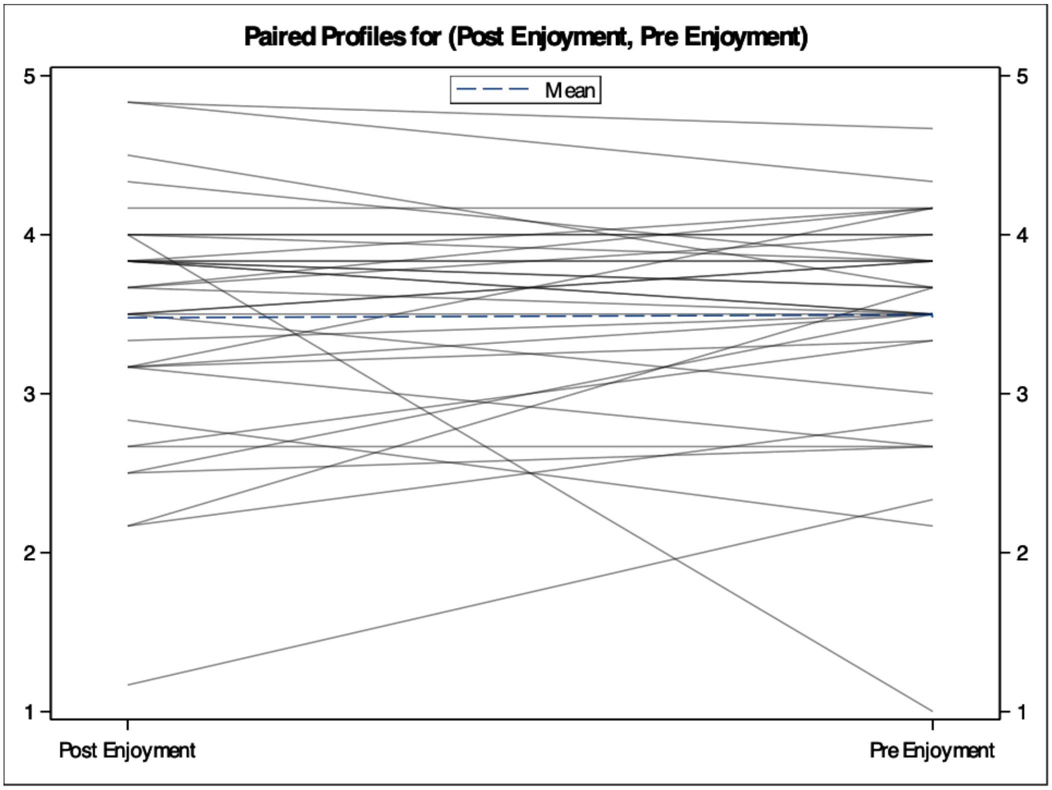
Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.847693	Pr < W	0.0002
Kolmogorov-Smirnov	D	0.156744	Pr > D	0.0239
Cramer-von Mises	W-Sq	0.197255	Pr > W-Sq	0.0052
Anderson-Darling	A-Sq	1.269209	Pr > A-Sq	<0.0050

N	Mean	Std Dev	Std Err	Minimum	Maximum
36	-0.0185	0.7344	0.1224	-1.5000	3.0000

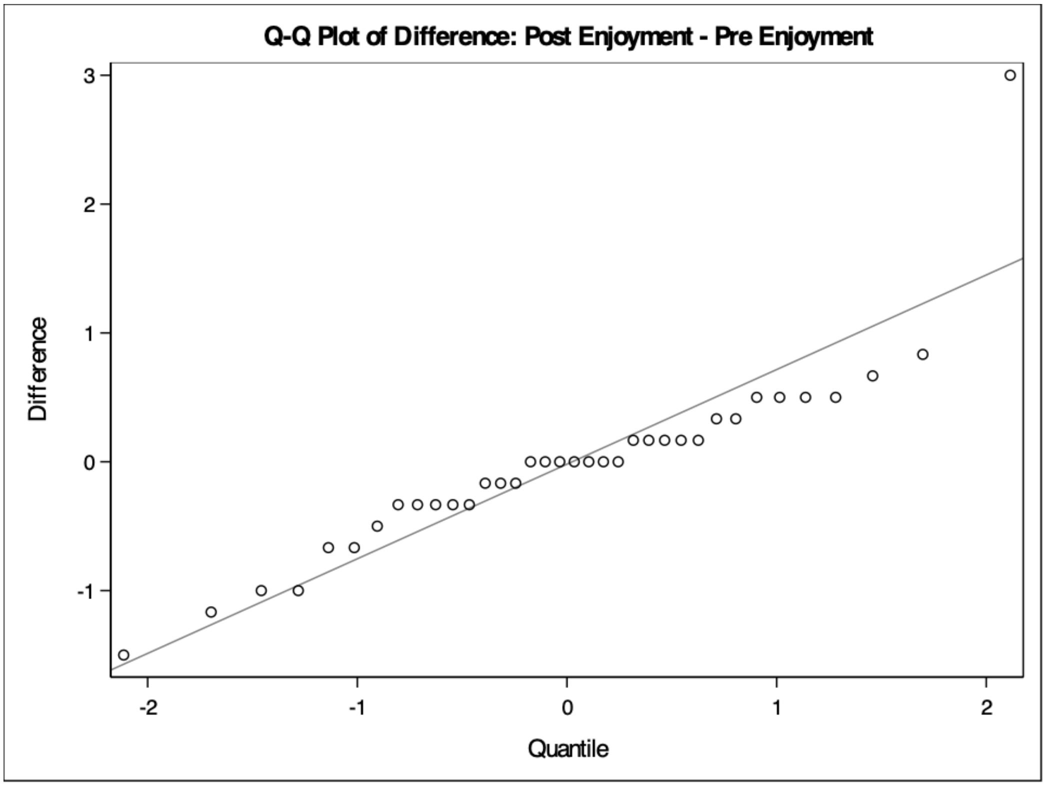
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-0.0185	-0.2670	0.2300	0.7344	0.5957	0.9580

DF	t Value	Pr > |t|
35	-0.15	0.8806




Paired Profiles for (Post Enjoyment, Pre Enjoyment)
Mean
Post Enjoyment
Pre Enjoyment
1
2
3
4
5


Agreement of Pre Enjoyment and Post Enjoyment
Mean
Pre Enjoyment
Post Enjoyment
1
2
3
4
5



Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	**-0.1513**	**Pr > |t|**	**0.8806**
Sign	**M**	**-0.5**	**Pr >= |M|**	**1.0000**
Signed Rank	**S**	**-18**	**Pr >= |S|**	**0.7037**

Post-survey T-test comparing Treatment and Control Groups

Control Group (0)

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.960099	Pr < W	0.2165
Kolmogorov-Smirnov	D	0.135516	Pr > D	0.0921
Cramer-von Mises	W-Sq	0.089861	Pr > W-Sq	0.1496
Anderson-Darling	A-Sq	0.556295	Pr > A-Sq	0.1444

Treatment Group (1)

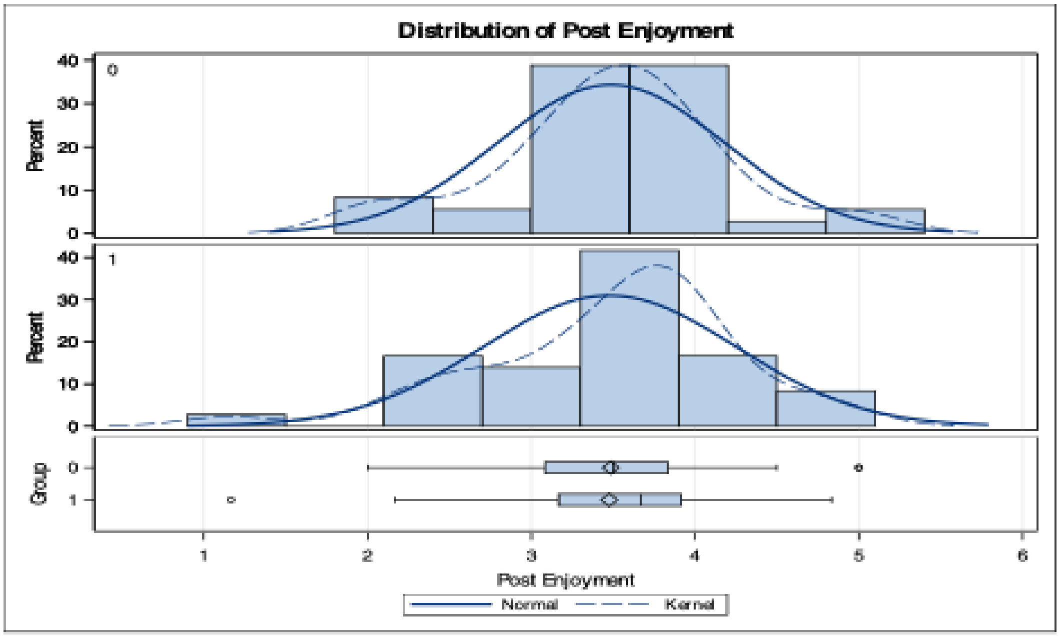
Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.945315	Pr < W	0.0745
Kolmogorov-Smirnov	D	0.150854	Pr > D	0.0375
Cramer-von Mises	W-Sq	0.142857	Pr > W-Sq	0.0286
Anderson-Darling	A-Sq	0.748458	Pr > A-Sq	0.0472

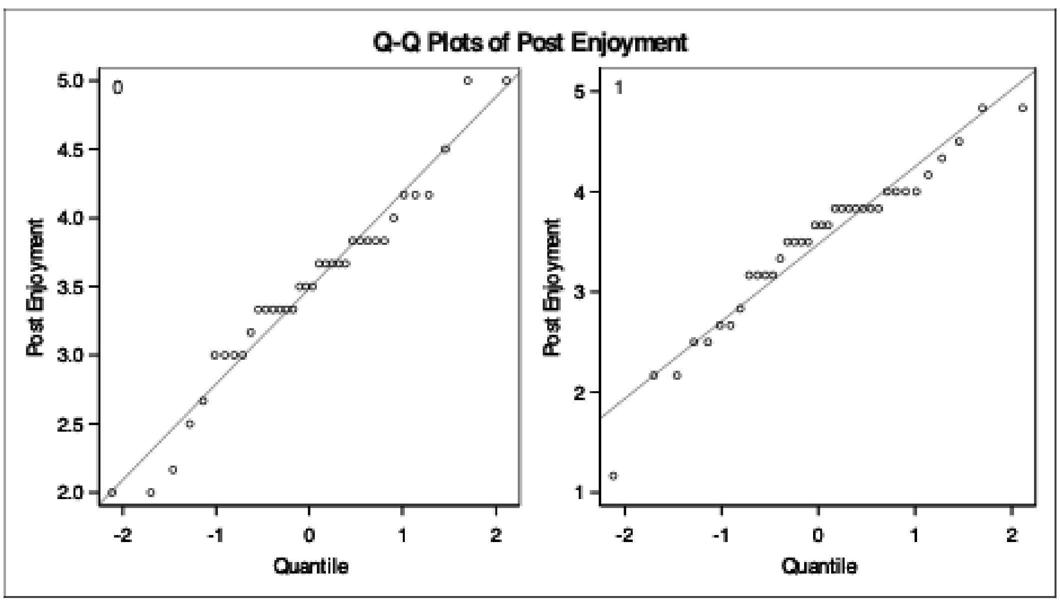
Group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		36	3.4861	0.6974	0.1162	2.0000	5.0000
1		36	3.4769	0.7717	0.1286	1.1667	4.8333
Diff (1-2)	Pooled		0.00926	0.7355	0.1733		
Diff (1-2)	Satterthwaite		0.00926		0.1733		

Group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3.4861	3.2502	3.7221	0.6974	0.5656	0.9097
1		3.4769	3.2158	3.7379	0.7717	0.6259	1.0066
Diff (1-2)	Pooled	0.00926	-0.3365	0.3550	0.7355	0.6312	0.8812
Diff (1-2)	Satterthwaite	0.00926	-0.3365	0.3551			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	70	0.05	0.9576
Satterthwaite	Unequal	69.294	0.05	0.9576

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	35	35	1.22	0.5523

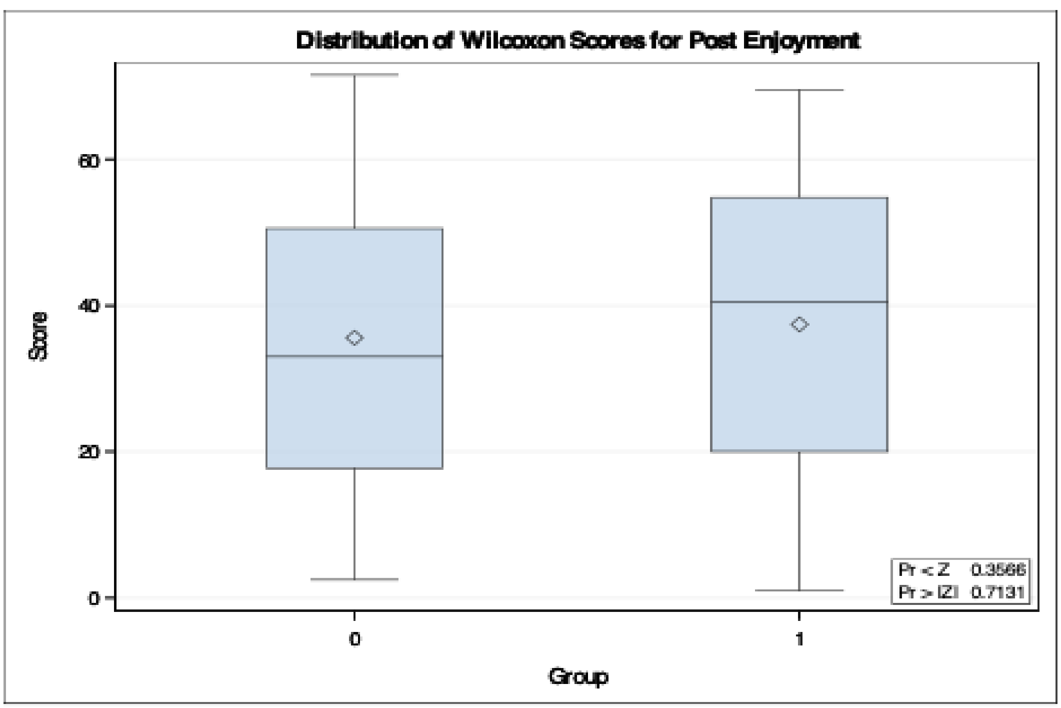




Wilcoxon Scores (Rank Sums) for Variable Post Enjoyment Classified by Variable Group					
Group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
0	36	1281.0	1314.0	88.392690	35.583333
1	36	1347.0	1314.0	88.392690	37.416667
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	Z	Pr < Z	Pr > |Z|	Pr < Z	Pr > |Z|
1281.000	-0.3677	0.3566	0.7131	0.3571	0.7142
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
0.1394	1	0.7089



Appendix J – Statistical Analysis for ATMI Subscale Motivation

Pre-survey T-test comparing Treatment and Control Groups

Control Group (0)

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.94824	**Pr < W**	0.0920
Kolmogorov-Smirnov	**D**	0.158949	**Pr > D**	0.0213
Cramer-von Mises	**W-Sq**	0.165824	**Pr > W-Sq**	0.0149
Anderson-Darling	**A-Sq**	0.894187	**Pr > A-Sq**	0.0212

Treatment Group (1)

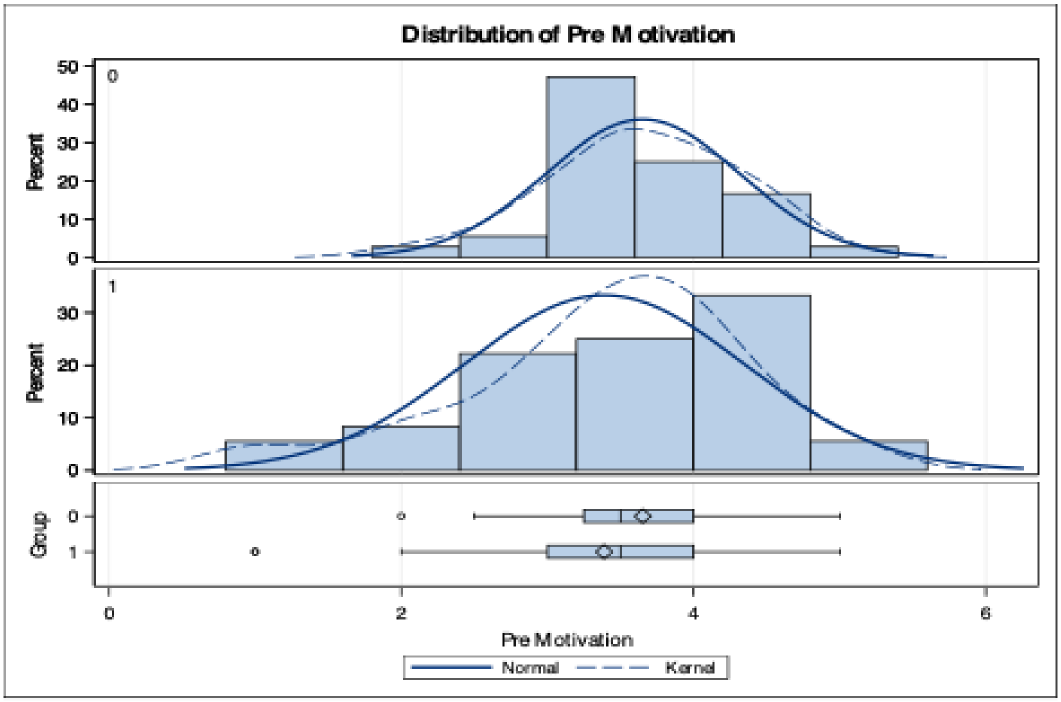
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.928516	**Pr < W**	0.0226
Kolmogorov-Smirnov	**D**	0.185103	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	0.175614	**Pr > W-Sq**	0.0100
Anderson-Darling	**A-Sq**	0.979436	**Pr > A-Sq**	0.0131

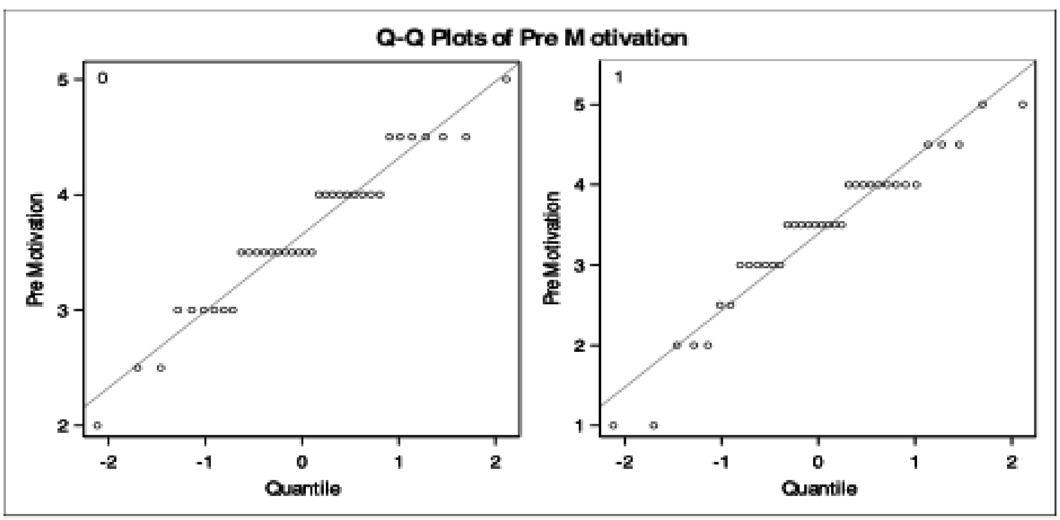
Group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		36	3.6528	0.6635	0.1106	2.0000	5.0000
1		36	3.3889	0.9570	0.1595	1.0000	5.0000
Diff (1-2)	**Pooled**		0.2639	0.8235	0.1941		
Diff (1-2)	**Satterthwaite**		0.2639		0.1941		

Group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3.6528	3.4283	3.8773	0.6635	0.5382	0.8655
1		3.3889	3.0651	3.7127	0.9570	0.7762	1.2484
Diff (1-2)	**Pooled**	0.2639	-0.1232	0.6510	0.8235	0.7068	0.9867
Diff (1-2)	**Satterthwaite**	0.2639	-0.1240	0.6518			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	70	1.36	0.1783
Satterthwaite	Unequal	62.334	1.36	0.1788

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	35	35	2.08	0.0333

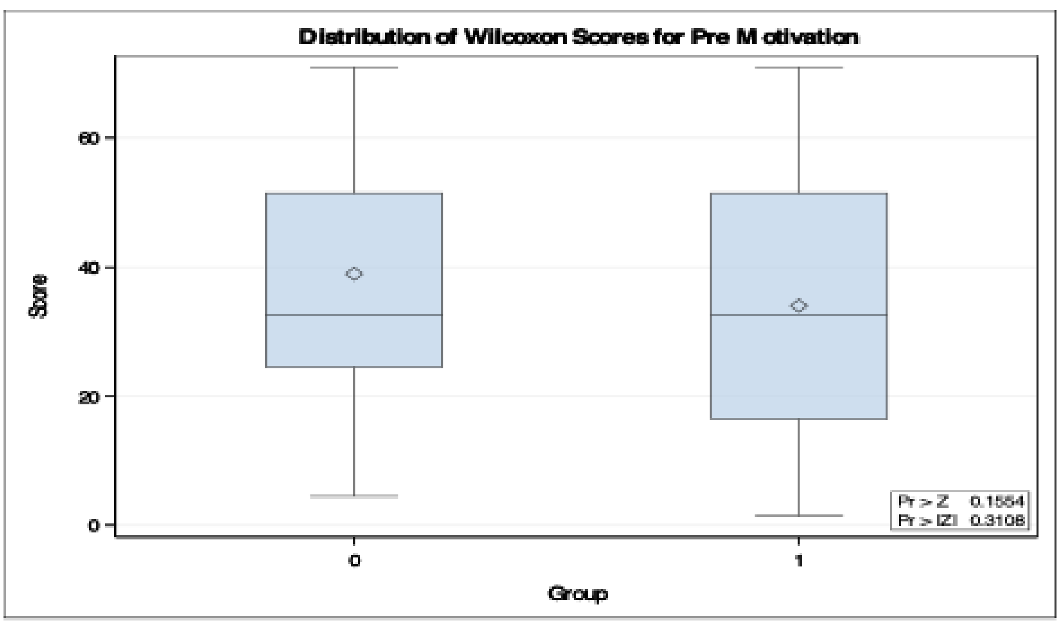




Wilcoxon Scores (Rank Sums) for Variable Pre Motivation Classified by Variable Group					
Group	**N**	**Sum of Scores**	**Expected Under H0**	**Std Dev Under H0**	**Mean Score**
0	36	1402.50	1314.0	86.821332	38.958333
1	36	1225.50	1314.0	86.821332	34.041667
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	**Z**	**Pr > Z**	**Pr > |Z|**	**Pr > Z**	**Pr > |Z|**
1402.500	1.0136	0.1554	0.3108	0.1571	0.3142
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	**DF**	**Pr > ChiSq**
1.0390	1	0.3080



ANCOVA results comparing Treatment and Control Groups

Class Level Information		
Class	Levels	Values
Group	2	0 1

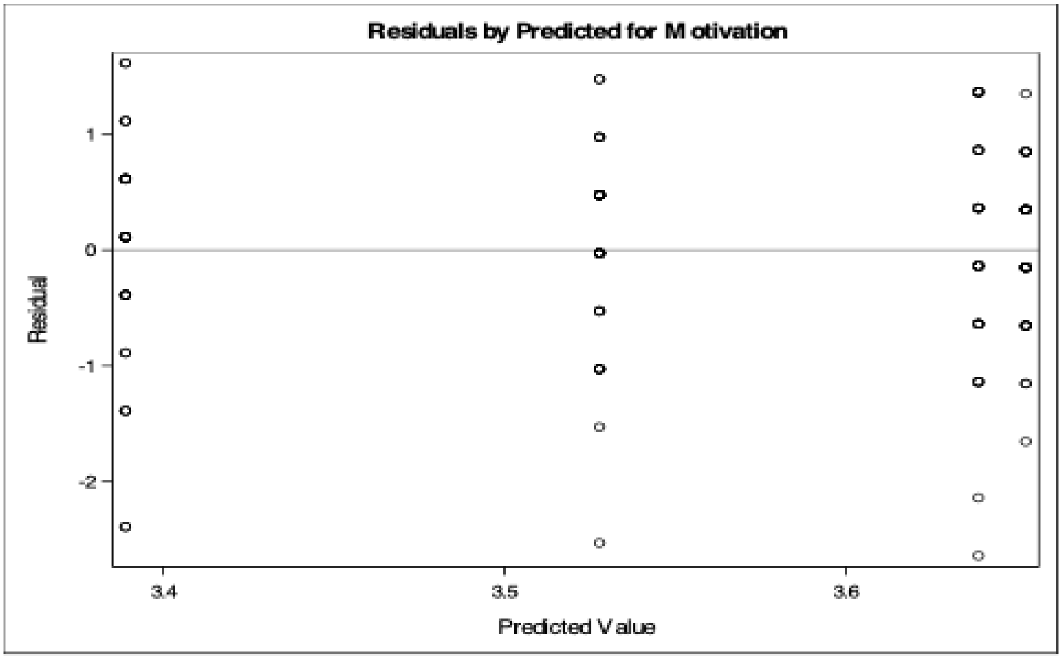
Number of Observations Read	297
Number of Observations Used	144

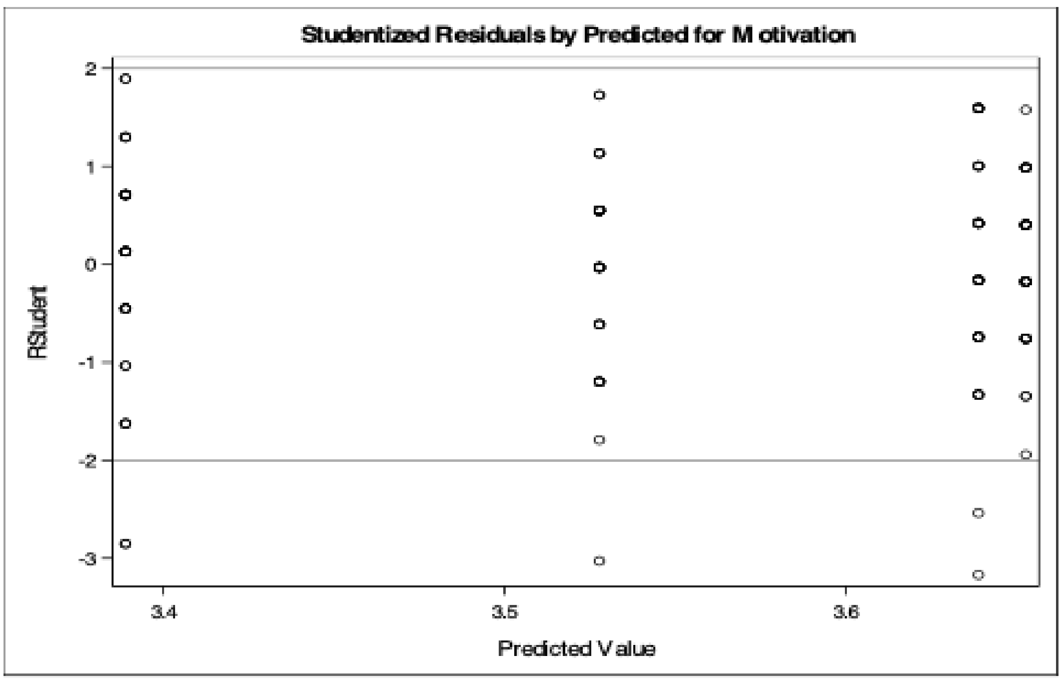
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1.6163194	0.5387731	0.71	0.5476
Error	140	106.2430556	0.7588790		
Corrected Total	143	107.8593750			

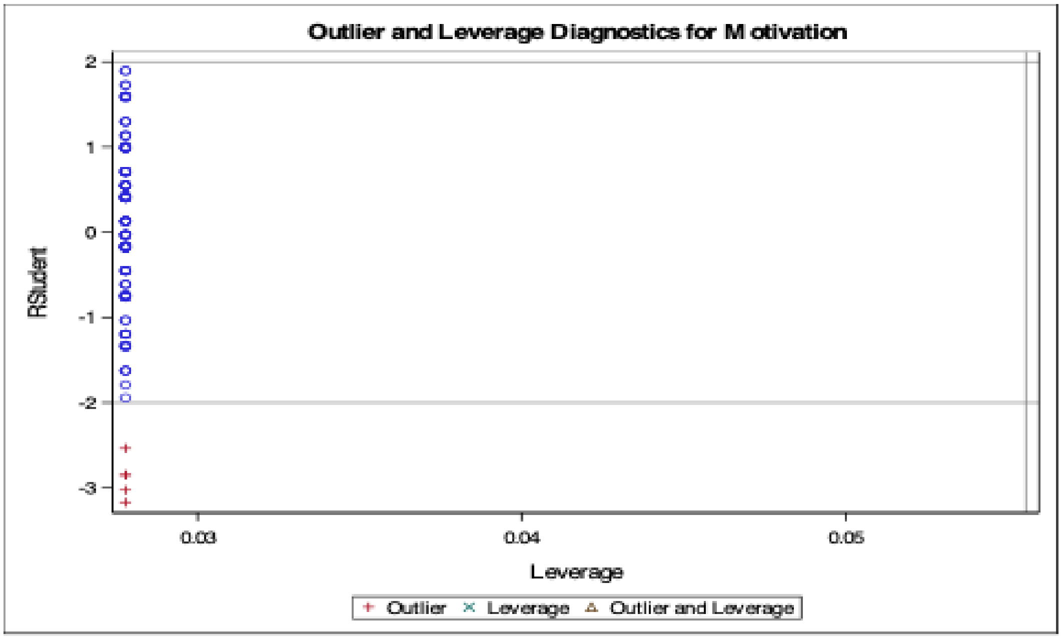
R-Square	Coeff Var	Root MSE	Motivation Mean
0.014985	24.52467	0.871137	3.552083

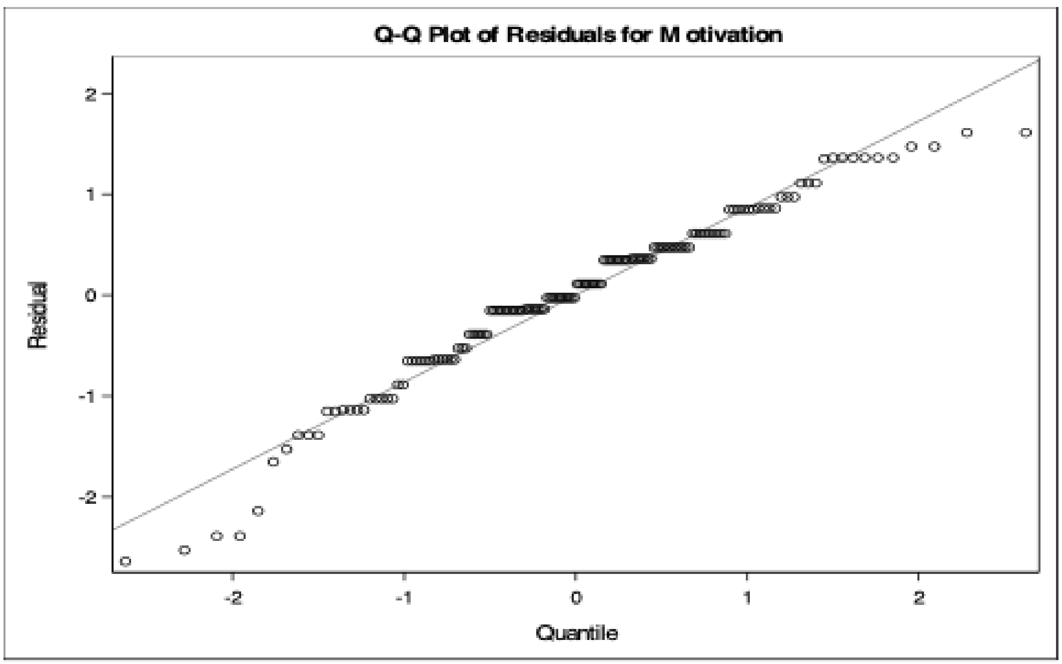
Source	DF	Type I SS	Mean Square	F Value	Pr > F
Group	1	1.26562500	1.26562500	1.67	0.1987
Pre-Post	1	0.14062500	0.14062500	0.19	0.6675
Pre-Post*Group	1	0.21006944	0.21006944	0.28	0.5996

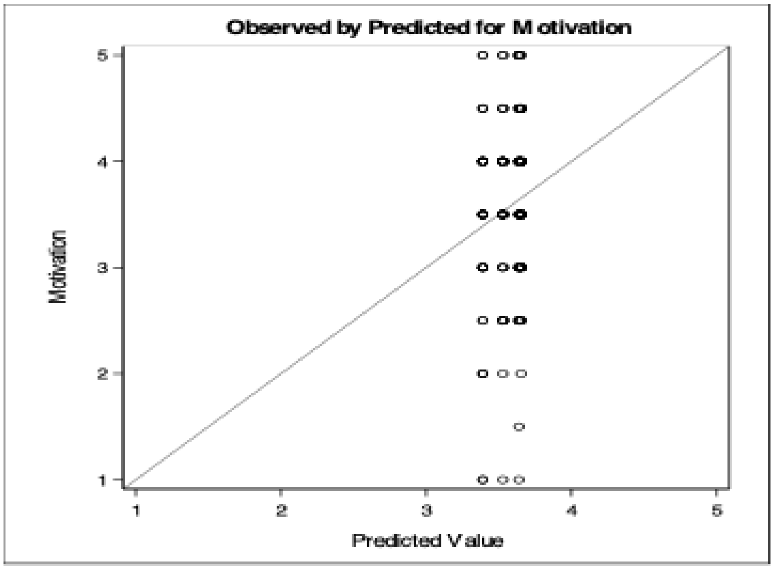
Source	DF	Type III SS	Mean Square	F Value	Pr > F
Group	1	1.25347222	1.25347222	1.65	0.2008
Pre-Post	1	0.14062500	0.14062500	0.19	0.6675
Pre-Post*Group	1	0.21006944	0.21006944	0.28	0.5996

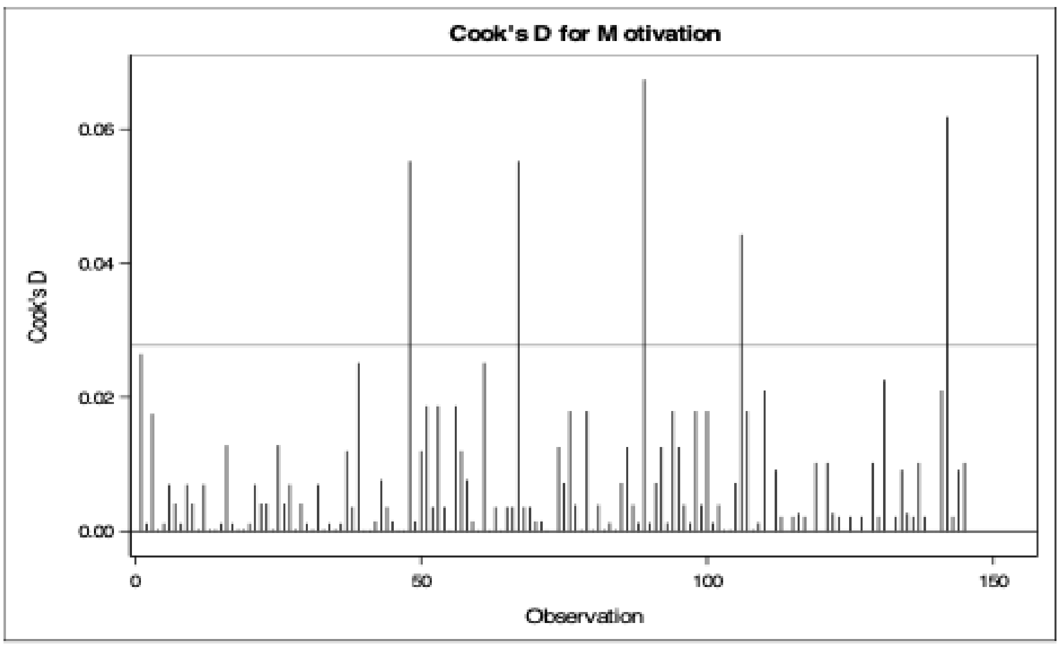

Residuals by Predicted for Motivation
Residual
Predicted Value

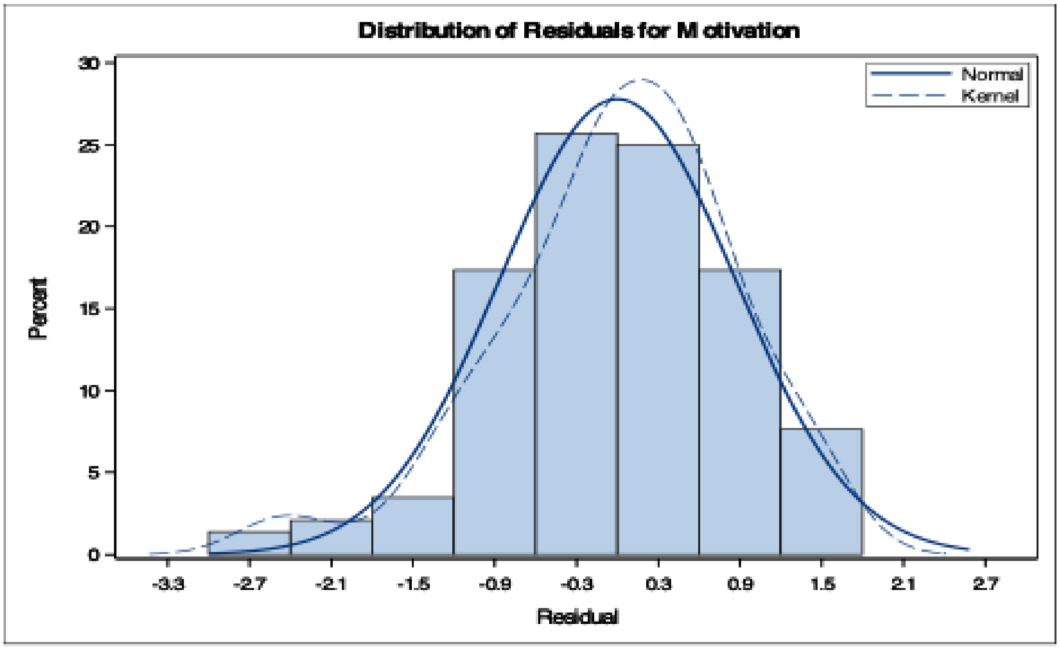

Studentized Residuals by Predicted for Motivation
RStudent
Predicted Value

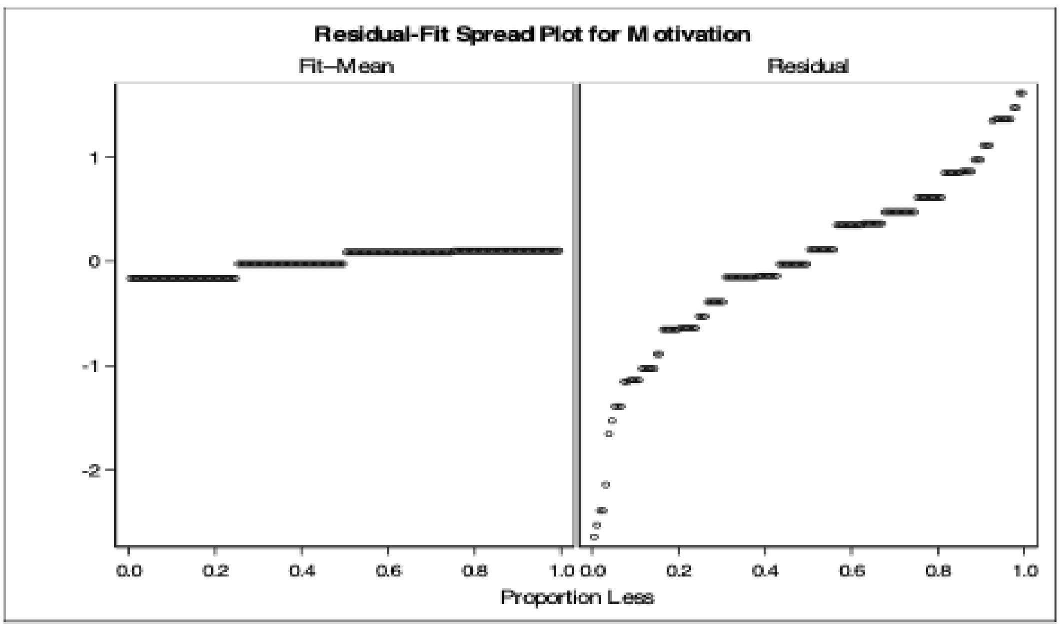

Outlier and Leverage Diagnostics for Motivation
RStudent
Leverage
0.03
0.04
0.05
Outlier
Leverage
Outlier and Leverage

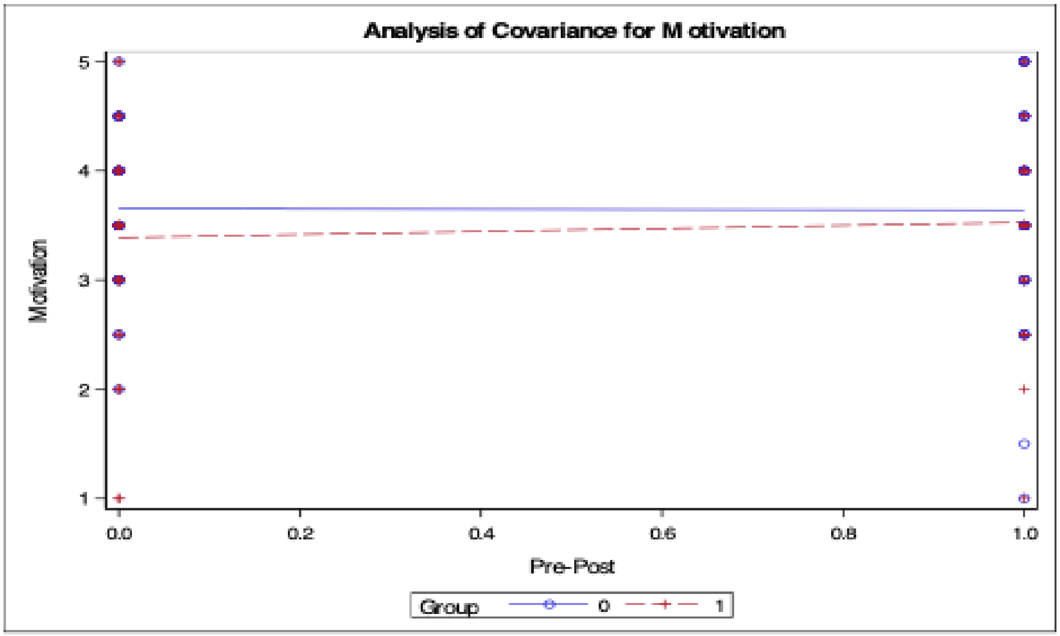

Q-Q Plot of Residuals for Motivation
Residual
Quantile


Observed by Predicted for Motivation
Motivation
Predicted Value
1
2
3
4
5

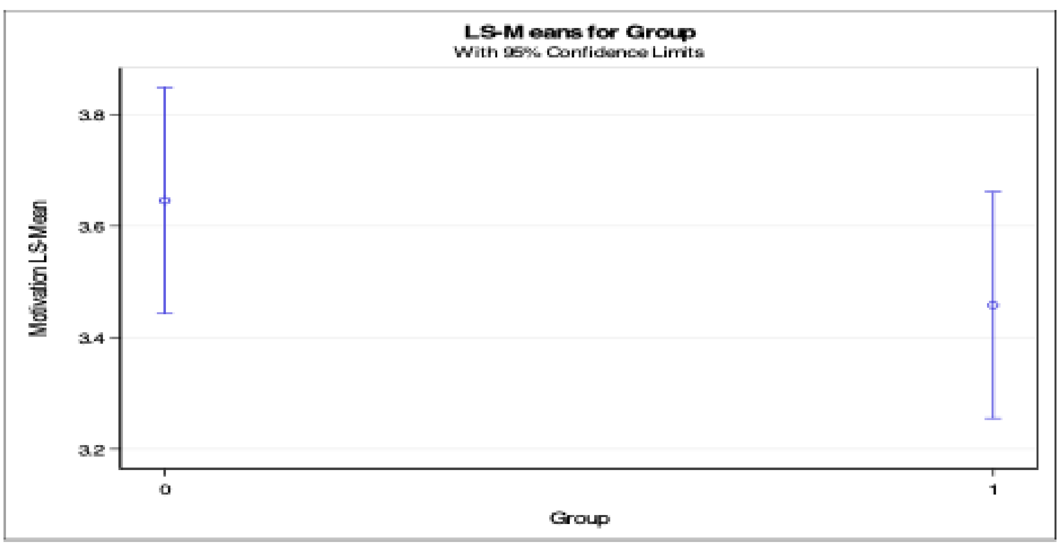

Cook's D for Motivation
Cook's D
0.00
0.02
0.04
0.06
Observation
0
50
100
150


Distribution of Residuals for Motivation
Normal
Kernel
Percent
30
25
20
15
10
5
0
-3.3
-2.7
-2.1
-1.5
-0.9
-0.3
0.3
0.9
1.5
2.1
2.7
Residual


Residual-Fit Spread Plot for Motivation
Fit–Mean
Residual
1
0
-1
-2
0.0
0.2
0.4
0.6
0.8
1.0
Proportion Less



Group	Motivation LSMEAN	H0:LSMean1=LSMean2 Pr > |t|
0	3.64583333	0.1987
1	3.45833333	



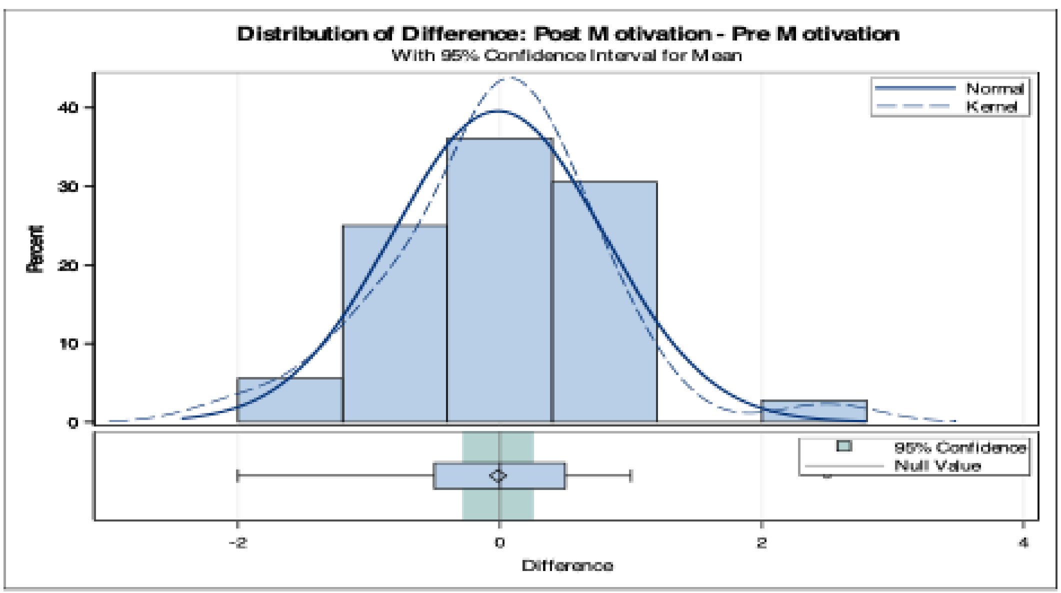
Pre-post Paired T-test Control Group

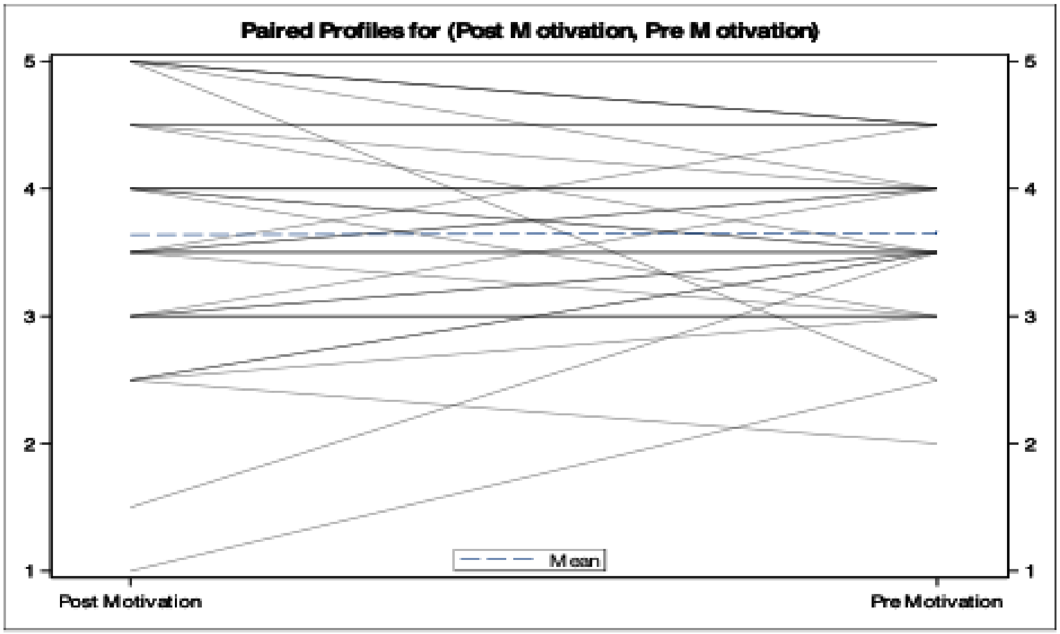
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.926713	**Pr < W**	0.0199
Kolmogorov-Smirnov	**D**	0.201318	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	0.217296	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	1.122812	**Pr > A-Sq**	0.0054

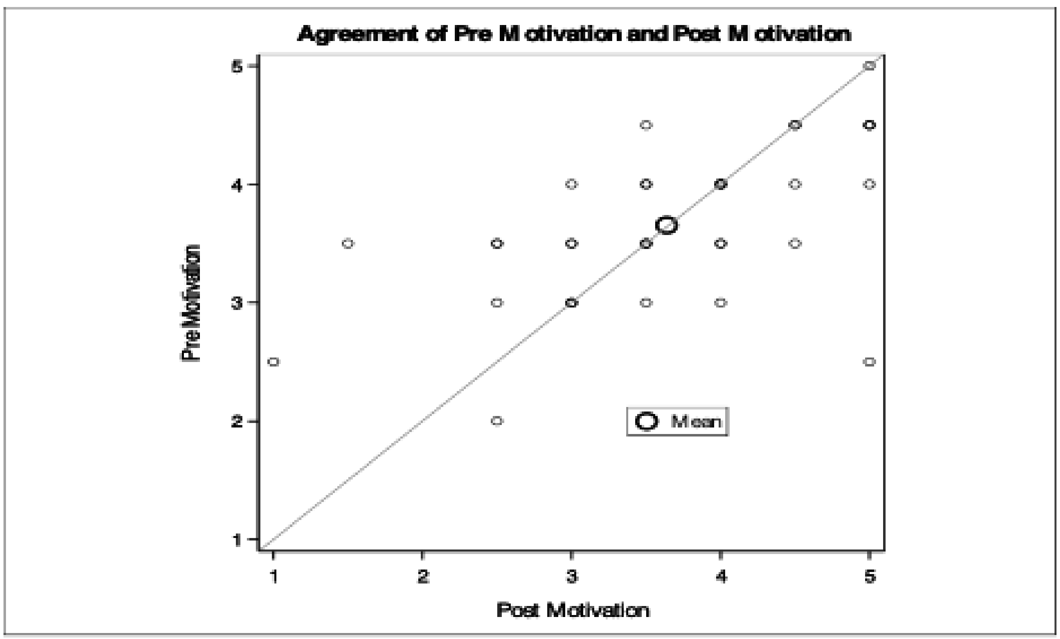
N	Mean	Std Dev	Std Err	Minimum	Maximum
36	-0.0139	0.8061	0.1344	-2.0000	2.5000

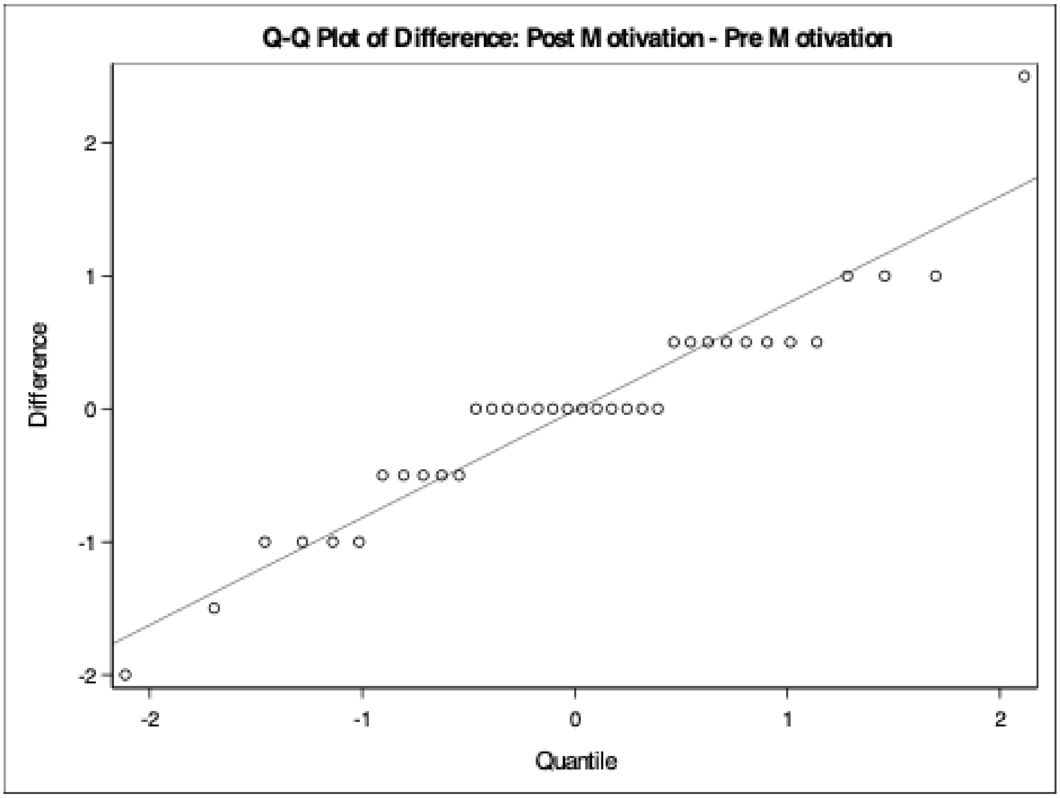
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-0.0139	-0.2866	0.2589	0.8061	0.6538	1.0515

DF	t Value	Pr > |t|
35	-0.10	0.9183




Paired Profiles for (Post M otivation, Pre M otivation)
5
4
3
2
1
Mean
Post Motivation
Pre Motivation


Agreement of Pre M otivation and Post M otivation
Pre Motivation
Mean
Post Motivation
1
2
3
4
5



Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	-0.10338	**Pr > |t|**	0.9183
Sign	**M**	0.5	**Pr >= |M|**	1.0000
Signed Rank	**S**	-8	**Pr >= |S|**	0.8093

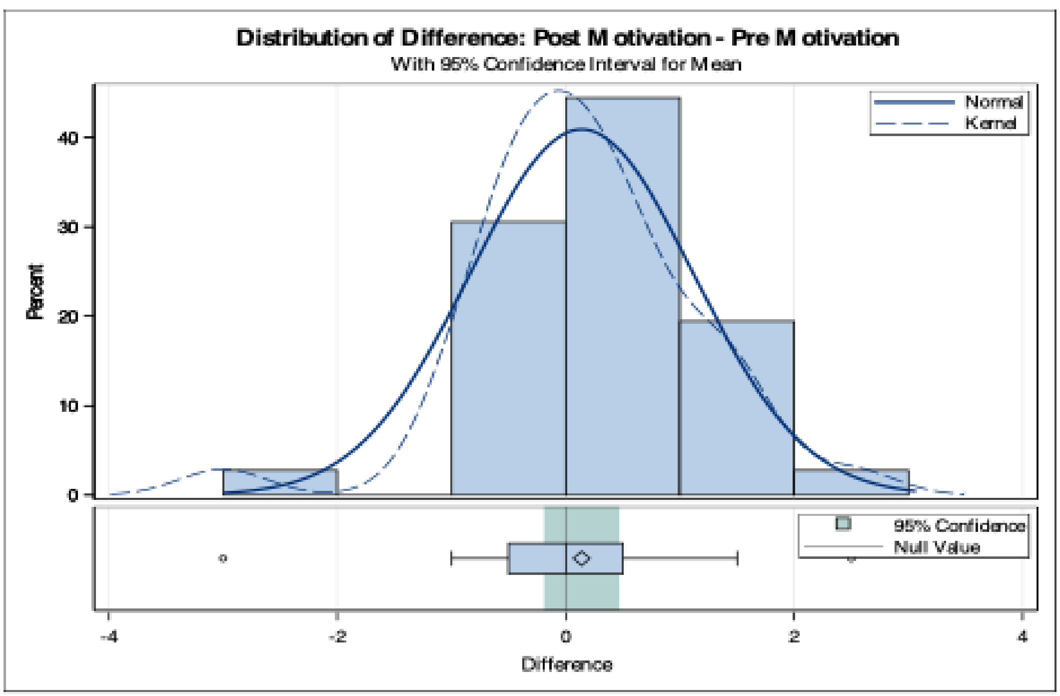
Pre-post Paired T-test Treatment Group

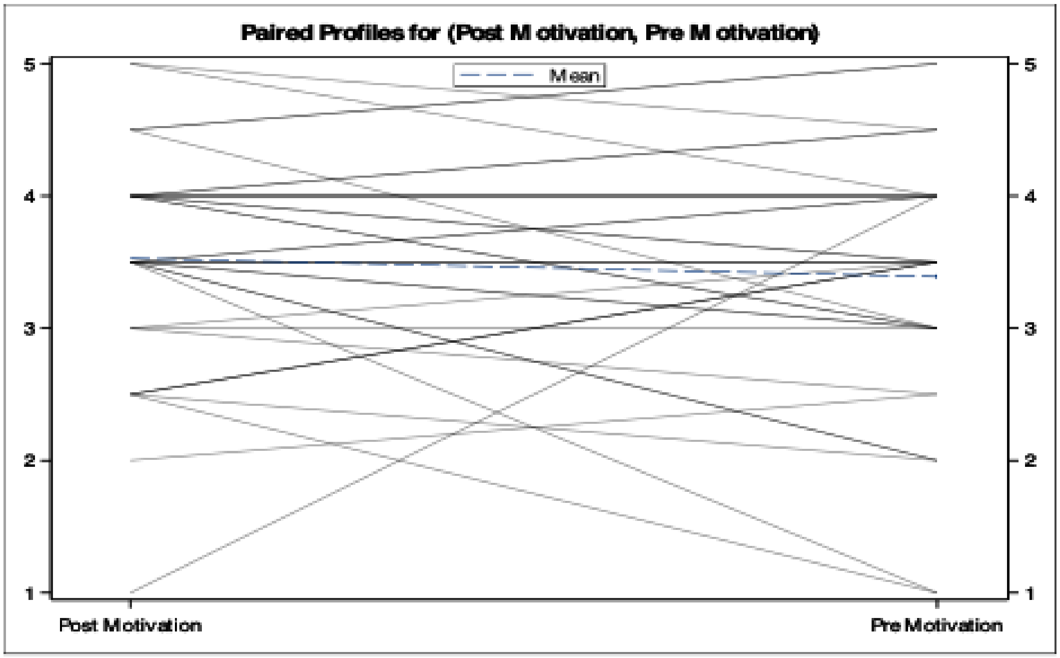
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.931854	**Pr < W**	0.0285
Kolmogorov-Smirnov	**D**	0.145142	**Pr > D**	0.0533
Cramer-von Mises	**W-Sq**	0.138561	**Pr > W-Sq**	0.0335
Anderson-Darling	**A-Sq**	0.845361	**Pr > A-Sq**	0.0268

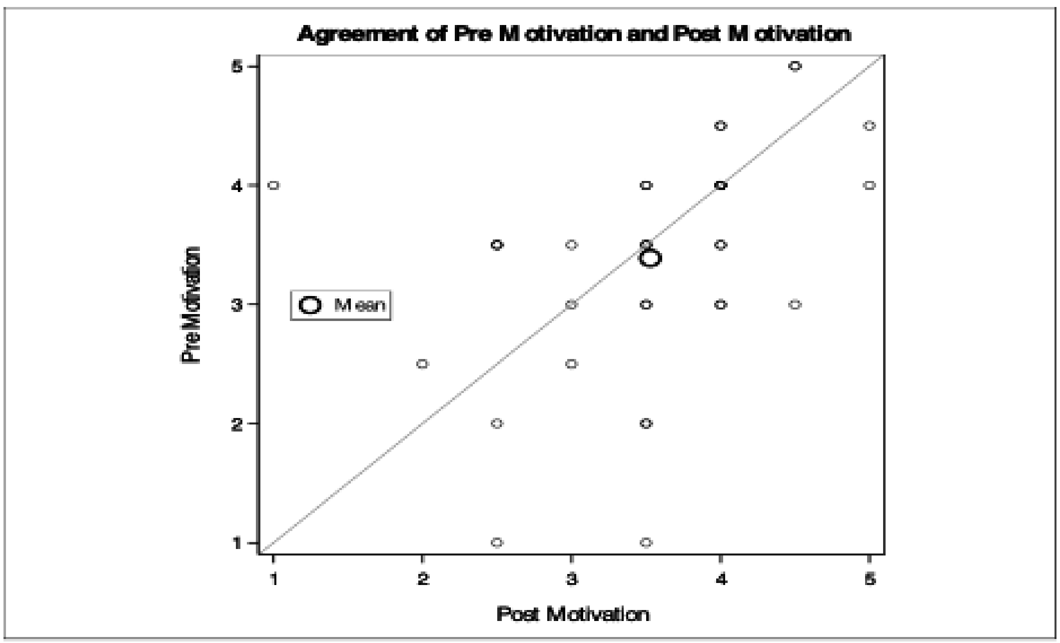
N	Mean	Std Dev	Std Err	Minimum	Maximum
36	0.1389	0.9755	0.1626	-3.0000	2.5000

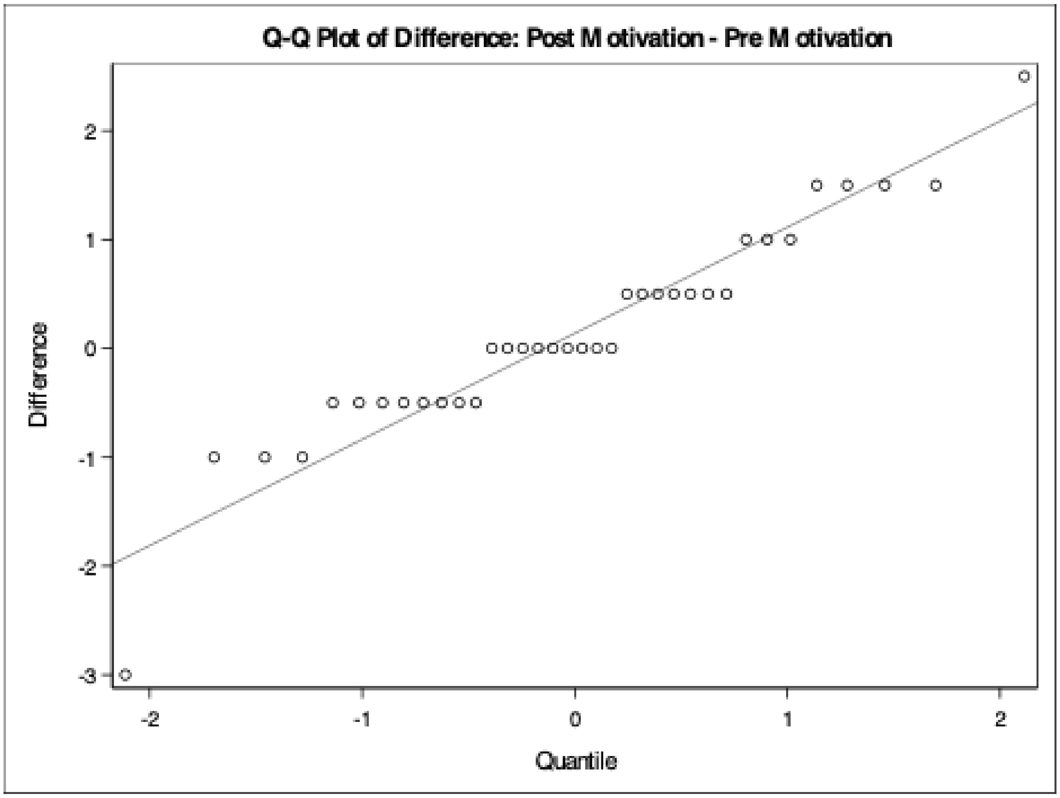
Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0.1389	-0.1912	0.4689	0.9755	0.7912	1.2725

DF	t Value	Pr > |t|
35	0.85	0.3988




Paired Profiles for (Post M otivation, Pre M otivation)
Mean
5
4
3
2
1
Post Motivation
Pre Motivation


Agreement of Pre M otivation and Post M otivation
Mean
Pre Motivation
Post Motivation
1
2
3
4
5



Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	0.854269	**Pr > |t|**	0.3988
Sign	**M**	1.5	**Pr >= |M|**	0.7011
Signed Rank	**S**	42.5	**Pr >= |S|**	0.3052

Post-survey T-test comparing Treatment and Control Groups

Control Group (0)

Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.938007	**Pr < W**	0.0441
Kolmogorov-Smirnov	**D**	0.114572	**Pr > D**	>0.1500
Cramer-von Mises	**W-Sq**	0.089362	**Pr > W-Sq**	0.1523
Anderson-Darling	**A-Sq**	0.635485	**Pr > A-Sq**	0.0922

Treatment Group (1)

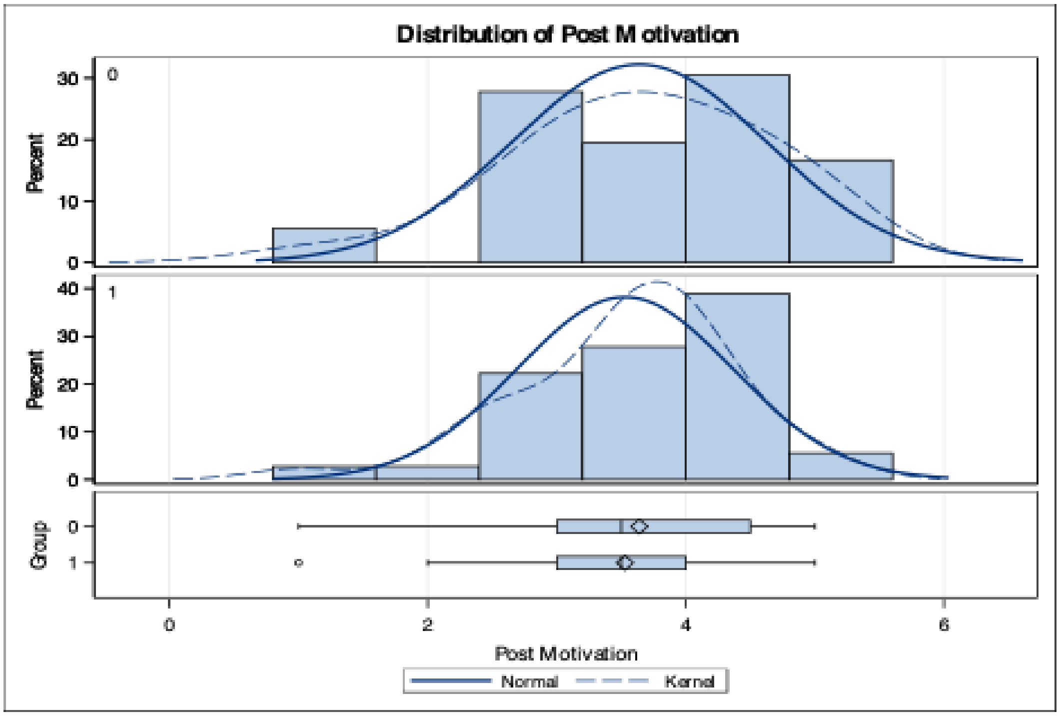
Tests for Normality				
Test	**Statistic**		**p Value**	
Shapiro-Wilk	**W**	0.925531	**Pr < W**	0.0184
Kolmogorov-Smirnov	**D**	0.208972	**Pr > D**	<0.0100
Cramer-von Mises	**W-Sq**	0.217411	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	1.118635	**Pr > A-Sq**	0.0056

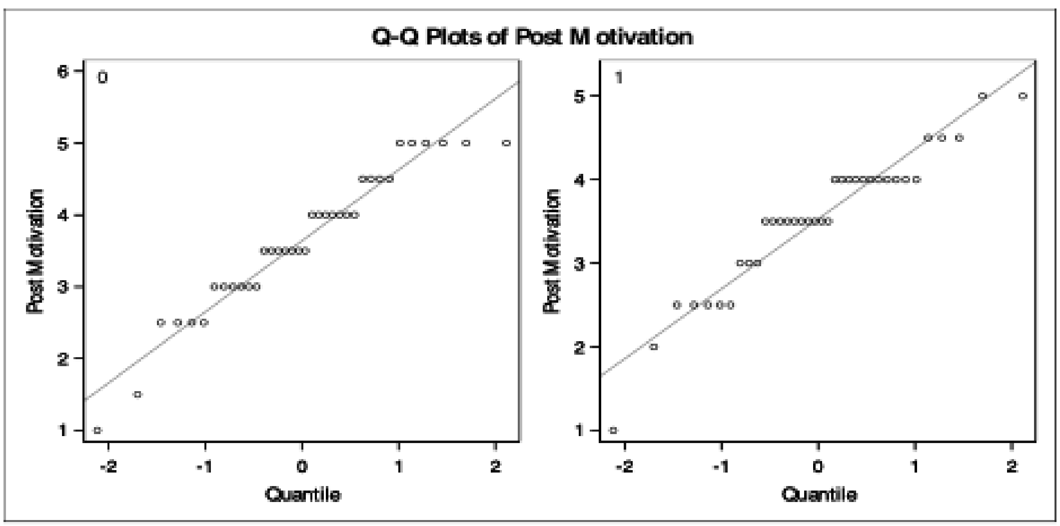
Group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0		36	3.6389	0.9900	0.1650	1.0000	5.0000
1		36	3.5278	0.8362	0.1394	1.0000	5.0000
Diff (1-2)	**Pooled**		0.1111	0.9163	0.2160		
Diff (1-2)	**Satterthwaite**		0.1111		0.2160		

Group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0		3.6389	3.3039	3.9739	0.9900	0.8030	1.2914
1		3.5278	3.2449	3.8107	0.8362	0.6782	1.0908
Diff (1-2)	**Pooled**	0.1111	-0.3197	0.5419	0.9163	0.7865	1.0980
Diff (1-2)	**Satterthwaite**	0.1111	-0.3199	0.5421			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	70	0.51	0.6086
Satterthwaite	Unequal	68.094	0.51	0.6086

Equality of Variances				
Method	**Num DF**	**Den DF**	**F Value**	**Pr > F**
Folded F	35	35	1.40	0.3223





Wilcoxon Scores (Rank Sums) for Variable Post Motivation Classified by Variable Group					
Group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
0	36	1360.0	1314.0	87.232085	37.777778
1	36	1268.0	1314.0	87.232085	35.222222
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
				t Approximation	
Statistic	Z	Pr > Z	Pr > |Z|	Pr > Z	Pr > |Z|
1360.000	0.5216	0.3010	0.6020	0.3018	0.6036
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
0.2781	1	0.5980