

**The Impact of Stability Ball Seating on
Behavior and Academic Achievement among Elementary
School Students in Grades 1 through 4.**

A Dissertation Presented for the
Doctor of Philosophy
Degree
University of Tennessee, Knoxville

Sheila R. Borders
December 2016

DEDICATION

I dedicate this research to my family, without whom I am nothing more than a wandering soul looking for purpose. First to my most dear, dedicated, and supportive husband, Alan Borders. I thank you for all you have done for me and my crazy ideas. In all you do, you exceed my expectations daily. I am only complete with you by my side and would never want it any other way. Thank you for being my world partner, reminding me that together we can accomplish anything. This is as much your dissertation as it is mine.

To my kids, Robby and Lilly Borders. I am the most blessed mother in the world to have such wonderful and understanding kids. This has been such a long journey that would not have been possible without your patience, understanding, laughs and hugs along the way. You will never know how much you each mean to me. Your jokes, your smiles, your support and your motivation are my lifeline that I will cherish for all eternity.

To my support network. There are oh so many of you to recognize and please know that you are remembered, each and every one of you even if you are not named here but a special thank you goes to Kim Juarez Hinton for her patience, understanding, pushing, prodding, and complete pestiferance. I truly value you as a support, a friend and wonderful human being.

Lastly, God. Without His support, love and forgiveness none of this would be possible.

Thank you all for your love, understanding and support.

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my advisor Dr. Clea McNeely for the continuous support of my Ph.D. study and related research, for her patience, motivation, and immense knowledge. Her guidance helped me in all the time of research and writing of this dissertation and I could not have imagined having a better advisor and mentor for my Ph.D. study.

Besides my advisor, I would like to thank the rest of my committee: Dr. Paul Erwin, Dr. Dawn Coe and Dr. Paul Terry, for their insightful comments and encouragement, but also for the hard question which inspired me to widen my research from various perspectives.

Thank you to UT Center for Health Policy and Services Research for partial funding of this research. Without this, the research would not have been possible.

My sincere thanks also goes to the administration of the University of Tennessee Extension. Without the support, both professionally and personally it would not be possible to conduct this research. The partial funding supplied through Family and Consumer Science Supplemental Grant program enabled the research to be implemented, the time allotted for the implementation, and the support and motivation are all without question the most valuable resources used to see this research through fruition.

ABSTRACT

This study was designed to answer three questions: 1) What is the impact of stability ball seating on the behavior of students in grades 1 through 4? 2) What is the impact of stability ball seating on the mathematics scores of students in grades 1 through 4? and 3) What is the impact of stability ball seating on the reading scores of students in grades 1 through 4? Research has shown that stability ball seating is linked to behavior and academic achievement among students diagnosed with attention deficit hyperactivity disorder (ADHD) and/or autism spectrum disorder (ASD). This study examines the effect of stability ball seating on behavior and academic achievement among students in mainstream classrooms in grades 1-4 who do not have a diagnosis of ADHD or ASD and was conducted in one elementary school in Lenoir City, Tennessee through a group randomized design. All classes in grades 1 through 4 agreed to participate. Data were collected from 136 students in first grade, 140 students in second grade, 141 in third grade and 80 in fourth grade, for an overall response rate of 99.9%. There were a total of 229 males and 268 females in this 16-week group randomized trial. Changes in classroom conduct, aggressive and attention deficit hyperactivity disorder behavior was monitored by teachers and identified through the Achenbach's Teacher Report Form which was completed as a pre and post intervention assessment with every student in the study. The Discovery Educational Assessment was also regularly implemented school-wide and scores for all students were monitored for any changes in academic achievement. Stability ball seating did not improve behavior or academic achievement in this sample. Any changes that occurred in student behavior or reading and math scores cannot be attributed to the stability ball seating but are probably the result of other influences such as teacher involvement, curricula, or other possible environmental factors excluding stability ball seating. It would not be recommended that teachers replace seating as an intervention in effort to improve behaviors as outlined in this research nor to improve reading or math scores.

TABLE OF CONTENTS

Chapter One: Introduction	1
Background: Education and Health	1
Strategies to Increase Early Academic Achievement.....	2
Seating Arrangement, Student Behavior and Academic Achievement	4
Research Questions	6
Chapter Two: Review of Literature	7
History	7
Dynamic seating for the student in general education.....	8
Stability ball seating with Autism Spectrum Disorder (ASD)	10
Stability ball seating with Attention Deficit Hyperactivity Disorder (ADHD).....	11
Chapter Three: Research Methodology.....	17
Problem and Purpose Overview	17
Research Setting.....	17
Study Design	21
Implementation of the Intervention.....	24
Data Collection Method.....	25
Measures	27
Teacher Report Form	28
Discovery Educational Assessment	36
Data Analysis Method	43
Power Analysis.....	47
Chapter Four: Results	54
Chapter Five: Summary, Conclusion, Discussion and Recommendations	64
Summary	64
Background	64

Methodology.....	65
Results.....	66
Conclusion	66
Discussion.....	67
Recommendation.....	67
Works Cited.....	69
Appendices.....	75
Vita.....	85

LIST OF TABLES

1.	Summary of research studied as it relates to stability ball seating in the classroom with students diagnosed with ADHD and ASD	16
2.	Adequate Yearly Progress Score (AYP) for Eaton Elementary School for years 2008-2012	19
3.	Distribution of students in control and intervention groups	23
4.	Questions from Teacher Report Form dealing with Conduct Problems, Aggressive Behavior and ADHD Behaviors	32
5.	Testing reliability coefficients from Cronbach's alpha for reading from Discovery Education Assessment	40
6.	Testing reliability coefficients from Cronbach's alpha for mathematics from Discovery Educational Assessment	40
7.	Comparison of proficient and/or advanced percentages for DEA and non-DEA schools in Metro Nashville, Tennessee for reading and mathematics	42
8.	Comparison of baseline data for gender, age, pretest ADHD, pretest Conduct, pretest Aggressive, January Math DEA, and January Reading DEA of control and intervention groups prior to intervention	43
9.	Comparison of pretest and posttest data on ADHD, conduct, aggressive behavior, math and reading scores.....	54
10.	Parameter estimates, standard errors and selected goodness-of-fit statistics for three random effects multilevel models describing the fitted relationship between conduct scores and ball intervention.....	56
11.	Parameter estimates, standard errors and selected goodness-of-fit statistics for three random effects multilevel models describing the fitted relationship between aggressive scores and ball intervention.....	58
12.	Parameter estimates, standard errors and selected goodness-of-fit statistics for three random effects multilevel models describing the fitted relationship between ADHD scores and ball intervention.....	59
13.	Parameter estimates, standard errors and selected goodness-of-fit statistics for three random effects multilevel models describing the fitted relationship between math scores and ball intervention.....	61

14.	Parameter estimates, standard errors and selected goodness-of-fit statistics for three random effects multilevel models describing the fitted relationship between reading scores and ball intervention	63
-----	---	----

LIST OF FIGURES

1. Effect of an additional 4 years of education on health outcomes.....	2
2. Standardized classroom seating76 with attached writing surface.....	4
3. Standardized classroom seating with detached writing surface.....	4
4. Diagram of pretest-posttest control group experimental design.....	22
5. Hierarchical breakdown of study population at Eaton Elementary School.....	44
6. Power Analysis of 29 classes with 25 students in each class.....	51

CHAPTER ONE

INTRODUCTION

Background: Education and Public Health

The education people receive can have an enormous impact on the rest of their lives. Higher-paying employment is a natural outcome of education, but that is not the only aspect of a person's life that will benefit from academic achievement. High education level is correlated with good health throughout out a person's life (Cutler and Lleras-Muney, 2007). Figure 1 demonstrates some of the benefits from an additional four years of education. These include: a 1.8 percent reduction in 5-year mortality; a 2.2 percent reduction in heart disease; a 1.3 percent reduction in diabetes; a 6.0 percent reduction in overall fair to poor health status; a 2.3 percent reduction in the number of sick days from work; an 11.0 percent reduction in smoking; a 7.0 percent reduction in heavy drinking; and, a 5 percent reduction in obesity (Cutler & Lleras-Muney, 2007).

Some studies assert that the link between health and education is due to higher-paying and rewarding employment, which leads to greater spending capacity, better housing, better access to health care, etc. Over time, education level has become a common indicator of socioeconomic success in many epidemiological studies (Winkleby, Jatulis, Frank, & Fortmann, 1992). The link, however, between education and health starts early in life. In re-evaluating data from a longitudinal study referred to as BCS70, Conti, Heckman, & Urzua (2010) observed that measures of health outcomes at age 30 are directly associated with cognitive and non-cognitive development established by the age of 10, a clear indication that early childhood education, conceived in a broad, holistic

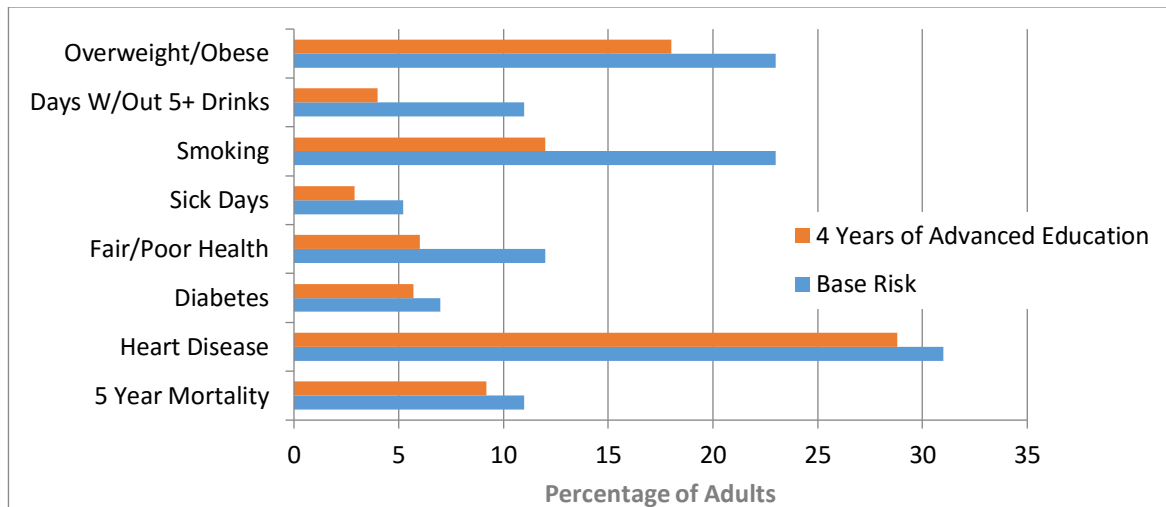


Figure 1: Effect of an additional 4 years of education on health outcomes (Cutler & Lleras-Muney, 2007)

way is an important factor throughout life. Lynch reports in his research from 2003 that not only does the link between education and health exist but it strengthens with age, the older have greater health and education links (Lynch, 2003).

Strategies to Increase Early Academic Achievement

Schools are well aware of the link between education and lifetime health outcomes. In effort to help students achieve the most while in their educational care, systems are willing to try new and innovative teaching techniques to not only help students but also increase test scores. With budget constraints ever increasing and expectations ever rising, school systems struggle to improve academic achievement for all students. With limited additional funding for individual classroom curriculum supplements, this task can be daunting for teachers. It is also stressful, as student achievement is increasingly linked to teachers' own professional success and funding for their school (No Child Left Behind (NCLB) Act of 2001, 2002).

Consequently, educators are examining many ways to help students grasp concepts and standards more effectively in order to maintain or increase standardized test scores. Different teaching styles, seating arrangements, classroom structure, and school-year calendar changes are all some of the strategies deployed to test theories regarding increased academic achievement. The learning environment is increasingly important in this endeavor as it is easier for teachers to manipulate and control the environment as they deem necessary than it is to address individual students' challenges, many of which derive from outside of school. Wannarka and Ruhl put it succinctly when they state, "...antecedent and structural events are of particular interest because manipulations of these may be considered naturalistic interventions that unobtrusively alter behaviors. Focusing on stimulus conditions rather than consequences may help teachers discretely to prevent problem behaviors before they occur and avoid utilizing unnecessarily intrusive interventions" (2008, p. 89).

Studies have been conducted on learning environment adaptations in an attempt to assist teachers improve their teaching effectiveness via environmental control (Capell, 2012; Cornell, 2002; Gur, Sackeim, & Gur, 1976; Guradino & Fullerton, 2010; Horwitz, 1979; Patton, Snell, Knight, & Gerken, 2001). The organization of the room, the arrangement of the chairs, and even the chairs themselves can affect the educational process just as do the practices the teacher implements and the information she conveys (Al-Eisa, Burgadda, & Melam, 2013; Bill, 2008; Cornell, 2002; Fedewa & Erwin, 2011; Gamache-Hulsmans, 2007; Guild, 2001; Guradino & Fullerton, 2010; Horwitz, 1979; Schilling & Schwartz, 2004).

In the next section, the literature regarding major strategies for environmental control will be briefly analyzed. Then, a more expansive account of the literature on seating and a specific focus on stability ball research, which forms the backdrop to my study, will be provided.

Seating Arrangements, Student Behavior, and Academic Achievement

Research has shown that seating has a link with academic achievement (Bill, 2008; Fedewa & Erwin, 2011; Gamache-Hulsmans, 2007; Schilling & Schwartz, 2004; Schilling, Washington, Billingsley, & Deitz, 2003; Wannarka & Ruhl, 2008; Weinstein, 1992; Wu, et al., 2012). Standard seating in most public school classrooms consists of a hard surface seat with a back made of the same material with metal legs. Sometimes these seats are independent of a desk, other times they have an attached writing surface. Please refer to figures 1 and 2 for clarification on standardized seating.



Figure 2: Standardized classroom seating with attached writing surface



Figure 3: Standardized classroom seating with detached writing surface

The most common seating alternative studied has been stability ball seating. In a couple of studies, the stability ball is referred to as a therapy ball, but the apparatus is still the same (Al-Eisa, Burgadda, & Melam, 2013; Schilling, Washington, Billingsley, & Deitz, 2003). The studies by Al-Eisa, Burgadda and Melam; Schilling, Washington, Billingsley and Deitz focus exclusively on students who were previously diagnosed with attention deficit hyperactivity disorder (ADHD) and/or autism spectrum disorder (ASD). Each of these studies discussed a link between the stability ball seat and improved classroom behavior and/or improved academic achievement. Students diagnosed with either ADHD or ASD were sometimes identified in the school system because of displaying behavior that repeatedly disrupts the classroom. Certain theories exist as to the reasoning behind the change in behavior. Gamache-Hulsman suggested the maturity of the student, teacher expectations and adjustment to classroom routine have possible links to the outcomes of the research (2007). Schilling and Schwartz admit to finding no concrete reason for the results but suggest investigating during a longer intervention as well as increasing the sample size to see if the impact remains (2004). Would these studies be replicable with students without such diagnoses or characteristics and have the same results? In other words, would seating intervention in the form of a stability ball improve the education attainment for all students in the general school population?

Research Questions

This dissertation answers the following research questions:

1. What is the impact of stability ball seating, used for 16 weeks in a regular classroom setting, on the conduct, aggressive and attention deficit hyperactivity disorder behavior of students in grades 1 through 4?
2. What is the impact of stability ball seating, used for 16 weeks in a regular classroom setting, on the mathematics scores of students in grades 1 through 4?
3. What is the impact of stability ball, used for 16 weeks in a regular classroom setting, seating on the reading scores of students in grades 1 through 4?

With funding ever the concern, schools need to be fiscally minded while still trying to provide students with a high-caliber learning experience so higher test scores will be rewarded with more funding. It is a cycle that seems to have no end. Small environmental changes within the classroom, such as stability ball seating, seem to offer the promise of higher academic success as well as being cost efficient. This study provides an evaluation of whether such a promise can be fulfilled for ordinary students.

CHAPTER TWO

REVIEW OF LITERATURE

History

The stability ball is often referred to as a “Swiss Ball” throughout research because though created in Italy, the ball was incorporated into physical therapy for the first time in Switzerland in the early 1960’s by Mary Quinton (Witt, 2001). Dr. Susan Klein-Vogelbach introduced the practice and helped to solidify the use of stability balls as a valid therapeutic aid (Carriere & Tanzenberger, 1998, pp. 1,2). Because of this addition to the physical therapy world, Swiss therapist Vlatka Zeller expanded the physical therapy clinic to the classroom (Witt, 2001). With the cooperation of an elementary school in Switzerland, she conducted an expansive test across Switzerland where five hypotheses were tested and supported (Witt, 2001):

1. Hyperactive children became calmer and would focus for longer periods of time
2. Other children could generally concentrate better
3. Handwriting skills improved with those exhibiting poor penmanship
4. Children often showed more understanding of subject matter being taught
5. Children were able to develop a better sense of organization

Since Zeller’s time there been a number of studies on stability ball seating in relation to classroom behavior and academic performance of students with special needs, specifically children with attention deficit hyperactivity disorder (ADHD) and autistic spectrum disorder (ASD) (Bagatell, Mirigliani, Patterson, Reyes, & Test, 2010;

Bill, 2008; Fedewa & Erwin, 2011; Gamache-Hulsmans, 2007, Schilling & Schwartz, 2004; Schilling, Washington, Billingsley, & Deitz, 2003; Witt, 2001)? It is unclear what the exact etiology linking improved behavior with stability ball seating is, but several studies have suggested fewer instructional interruptions as the mechanism between stability ball seating and reducing behavioral disruptions (Bill, 2008; Schilling, Washington, Billingsley, & Deitz, 2003; Wannarka & Ruhl, 2008). For example, Gettinger & Seibert (2002, 2003) argue that students seated on stability balls are able to have a subtle but effective energy outlet, causing less disruption during class, which, in turn, increases the overall instruction time for which they are exposed. In the next few sections, an analysis of the literature on the effectiveness of stability ball seating on Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD) will be provided.

Dynamic seating for the student in general education

This research will focus on the potential benefits of stability ball seating on the general classroom participant, the student without a clinical psychiatric diagnosis or appearance of symptoms signaling a possible future diagnosis of ADHD and/or ASD. Some research exists that can make the bridge to support this step from ADHD and ASD to the general student. Research indicates that some benefits of dynamic seating include improved behavior and academic achievement. In research by Sines, stability ball seating implemented in a classroom of second-grade students was a major component in the increased attention and time on task (Sines, 2009). The increased time on task also speaks to the decrease in negative behavior issues. However, seating

was only one component of the study, and it is not clear how much dynamic seating contributed to the overall results.

In their 2014 study, Budd and Rehling explore the possible biological bases for the success of utilizing stability ball seating as a tool for expanding educational success. Their hypothesis is that there is a plasticity in the central nervous system that can be positively influenced by stimulating sensory experiences (Budd & Rehling, 2014). These experiences can enrich and alter brain structure and motor performance. Through this alteration, changes are more profound when students interact more with their environments. The stability ball presents a condition for the students to move about the classroom, work in specific centers, and interact with their environments. The merger of sensory interaction with the learning process adds a layer of involvement to the student's development and therefore supports engagement and development in the learning environment (Budd & Rehling, 2014). This may occur due to the sensory processing theory as explained by Ivory (2011). According to this theory, sensations are used to support attention and behavior. With more sensations within the seating and work area, attention is drawn to that area, thus increasing time on task. The gains in terms of time on task will lead to increased productivity in the classroom through increased proprioceptive input and vestibular input (Ivory, 2011).

Added sensation in learning affects the vestibular system, which regulates head movement and balance, and the proprioceptive system, which registers body positioning, thereby inhibiting the over-response to other sensations that often lead to on-task disruptions or unfavorable behavior. Ivory's hypothesis is that this added

sensation therefore increases behavioral organization and helps students increase their ability to focus on the task at hand (Ivory, 2011).

John Medina, a neurobiologist, elaborates that incorporating movement for effective learning is the optimal environment because of the way human beings evolved as a species (Medina, 2009). According to Medina there is a direct link between movement and mental activity. Movement increases blood flow, which increases oxygen to the brain. With increased oxygen to the brain mental acuity increases. Therefore, incorporating movement into the classroom can increase attention, increase focused learning, and could increase academic achievement (Medina, 2009).

Stability ball seating with Autism Spectrum Disorder (ASD)

Students' struggles while coping with ASD in the classroom are difficult to pinpoint, as ASD is such a broad spectrum with varied symptoms. Most often, however, one commonality is present: these students all exhibit impediments to social engagement. These manifest through challenges students face relating to education faculty as well as to friends and peers (Whalon & Hart, 2010). Interactions with education faculty are required to learn; without that connection or engagement, key points can often be missed, resulting in lower academic achievement (Whalon & Hart, 2010).

Stability ball seating has been linked to increased time on task for ASD students (Schilling & Schwartz, 2004). Through the dynamic seating of the stability ball, ASD students are able to focus on immediate surroundings and engage in the lessons longer than when seated in traditional seating (Schilling & Schwarz, 2004). Schilling has

published several articles on stability ball seating's impact on students with ASD. In one important study, Schilling and Schwartz theorized that sensory modulation strategies that engage the central nervous system create an alertness in students with ASD (Schilling & Schwarz, 2004). They worked with four male students, ranging in ages from 3 years 11 months to 4 years 2 months. Each student had been clinically diagnosed with ASD. Implementation of the seating alternative was specifically designed to accommodate each student. For example, one student used stability ball seating at the reading station for 10 minutes at a time and only after recess. Another student used it for 10 minutes only right after lunchtime with group art time. A third student could use the stability ball seating during circle time when the teacher read picture books and could only use it for 5 minutes. The fourth student could use the stability ball seating from 5 to 10 minutes right before recess while doing art activities (Schilling & Schwarz, 2004). With such specific time slots and durations researchers were able to collect data based on two criteria, correlating time spent on the stability ball with engaged behavior. Results indicate that all four students increased positive behavior and engaged behavior while seated on the stability ball (Schilling & Schwarz, 2004). With these results, it appears stability balls may have a positive impact for students with ASD. However, with only four students with 5 to 10 minutes of exposure, greater scope within the research is needed before a broad conclusion can be drawn.

Stability ball seating with Attention Deficit Hyperactivity Disorder (ADHD)

Students diagnosed with ADHD often struggle with academic achievement for various reasons. Common characteristics of students with ADHD that interfere with academic engagement include lack of impulse control, limited attention capacity, and

hyperactivity (Daley & Birchwood, 2009; Junod, DuPaul, Jitendra, Volpe, & Cleary, 2006). The foundation of learning during the preschool years is often jeopardized, resulting in decreased reading skills, a common problem with students with ADHD (Daley & Birchwood, 2009). Without this primal foundation for reading, all other subjects tend to suffer as well, resulting in students often becoming labeled as having learning disabilities (DuPaul, 2006).

Stability balls have helped students improve their focus, engagement, and academic achievement when otherwise dealing with ADHD (Bill, 2008) (Schilling & Schwartz, 2004). Bill (2008) placed stability balls in a classroom of students diagnosed with ADHD and found that students increased the proportion of time they spent on-task during instruction time. Schilling and Schwartz (2004) used video and audio recordings to evaluate students with ASD who were seated on stability balls for a shorter period of time during the day. Student engagement increased and behavior problems decreased during that time. There are a couple of possible mechanisms that could be responsible for these improvements. One such mechanism is the fact that students are able to “fidget” and release energy without disturbing other students or disrupt instruction time. Also, the engagement needed to balance on the ball may translate into overall engagement, resulting in improved academic achievement.

A common finding throughout stability ball research with ADHD students is that while they are seated on stability balls, movement and fidgeting are much less noticeable and disruptive. This consequence, in turn, requires less teacher intervention for disruptive behavior, resulting in increased instruction time. The movement on stability ball seating does not seem to interfere with surrounding students; therefore,

teachers can overlook the movement and continue teaching (Bill, 2008; Fedewa & Erwin, 2011; Harlacher, Roberts, & Merrell, 2006).

Schilling, et al, also researched the impact of stability ball seating on students diagnosed with ADHD. The students were members of a fourth-grade class consisting of 24 students, but only three were involved in the intervention. The three students diagnosed with ADHD used stability ball seating during their Language Arts class (Schilling, Washington, Billingsley, & Deitz, 2003). The intervention lasted for twelve weeks. The findings support the use of stability ball seating for the improvement of in-seat behavior and legible word productivity. Each of the three students increased seating intervals on average by 60 percent (Schilling, Washington, Billingsley, & Deitz, 2003). The limitations of this study, however, are fairly obvious: it included only three students and only a limited exposure to the intervention--one class period—throughout the day. There were also no validated assessments to measure academic achievement, though teachers expressed their belief that academic improvement had occurred (Schilling, Washington, Billingsley, & Deitz, 2003).

Fedewa and Erwin published similar research regarding a stability ball intervention with eight students in 4th and 5th grades diagnosed with ADHD. Trained student observers visited the classroom and monitored student behavior for 30 minutes, three times a week, to determine the effect of the alternative seating (Fedewa & Erwin, 2011). Observers categorized their data into a series of codes describing student behavior once every 30 minutes. Observers documented the duration of in-seat and on-task behaviors displayed by the students during a twelve-week intervention period. The study's conclusions support the effectiveness of use of stability ball seating with

students diagnosed with ADHD. Prior to the intervention, the eight students on average were seated 45 percent of the time and spent only 10 percent of their time on task (Fedewa & Erwin, 2011). After the intervention, Fedewa and Erwin (2011) found the same eight students were seated 94 percent of the time and spent approximately 80 percent of their time on task. The study included no control group for comparison or correction for any learning that would have taken place without the stability ball intervention. And the population of the study included only eight students, all diagnosed with ADHD and the disruption or expectations of the visitors may be a limitation of this intervention.

Capell worked with eighteen 1st grade students with ADHD and compared their math scores on tests prior to and after they utilized stability balls (2012). There was a slight increase in math scores in this research, but with no control group, it is not clear that the difference is due to the stability ball intervention as she states in her research. The eighteen students took math assessments, the intervention was implemented, and the students were immediately given the same test afterwards, all in a single class day. (Capell, 2012).

Limited research exists identifying the impact on stability ball seating on students not identified with ASD or ADHD diagnoses (Al-Eisa, Burgadda, & Melam, 2013; Amstad, Bachlin, & von Arx, 1992). As stated earlier, with several articles suggesting positive impacts of stability ball seating with students with ASD and ADHD and with research showing the positive impact of varied seating for students in the classroom in relation to academic achievement a logical question would be, does the stability ball seating impact students not diagnosed with ADHD or ASD in a positive manner as well?

If a stability ball seat is beneficial for students with ADHD or ASD one could think the same might true for students who are not dealing with these diagnoses. The research presented in this dissertation will explore whether stability ball seating is associated with classroom behavior and academic achievement in a general population of students in grades 1 through 4 in a school of Pre-K through 4th grade. Pre-K and kindergarten teachers asked to not be a part of the research and those requests were observed.

This reasonable proposition—that utilizing stability balls in place of traditional seating allows students to have an effective energy outlet, allowing for more overall instruction time— leads to a natural further question that forms the basis of this dissertation. Might the proposition also prove true for students in the general population, as opposed to students who suffer from particular disabilities that hamper learning? Surprisingly perhaps, no peer-reviewed research has been published to indicate that there is any correlation between the use of stability balls and improvements in exposure to instruction time. Perhaps the mechanism that is the root of improved behavior is the fact that students are made to focus on their immediate surroundings, maintain balance and therefore are less distracted by outside stimuli as hypothesized by Schilling and Schwartz (2004). This review of literature shows that a preponderance of research indicates that investigating seating alternatives in general classroom settings is an appropriate and potentially fruitful endeavor. The research of Winkleby, et al. indicates that the best possible window for affecting the future health of adults is by examining educational experiences of youth prior to age 10; thus, elementary school students should be a prime focus for analyzing the effectiveness of alternative seating.

Table 1

Summary of research studied as it relates to stability ball seating in the classroom with students diagnosed with ADHD and ASD.

Source	Sample	ADHD	ASD	Academics	Behavior	Other
Al-Eisa, Burgadda & Melam, 2013	40					Improved comfort
Bagatell, Mirigliani, Patterson, Reves, & Test 2010	6		X	Decrease	Decrease	
Bill, 2008	12	X		Increase	Increase	
Capell, 2012	18			Increase		
Fedewa & Erwin, 2011	8	X		Increase	Increase	
Harlacher, Roberts, & Merrell, 2006	30*	X		Increase	Increase	
Schilling & Schwartz, 2004	4		X	Increase	Increase	
Schilling, Washington, Billingsley, & Deitz, 2003	3	X		Increase	Increase	

**simply states 1 class*

CHAPTER THREE

RESEARCH AND METHODOLOGY

Problem and Purpose Overview

The purpose for this study is to determine the impact of stability ball seating on behavior and math and reading scores among students in grades 1 through 4. For the purposes of this research, behavior is a culmination of student classroom habits scored via the Teacher Report Form measuring the specific syndromes of attention deficit hyperactivity disorder, conduct, and aggressiveness.

Research setting

This research took place in one school, Eaton Elementary School (Eaton) in Lenoir City, Tennessee. Eaton is one of nine schools in the Loudon County school district and one of three elementary schools in Lenoir City, Tennessee. One reason for selecting Eaton is because school administrators were open to the idea of a classroom intervention, in part due to the author's previous working relationship with this school and the teachers' willingness to cooperate. Students were in self-contained classrooms—with the exception of lunch or recess, students remain in the same classroom most all of the day. This was beneficial for this research as the students' teachers are the same and educational environment remained relatively constant. It also alleviated any obstacles in terms of transporting stability balls between classrooms.

In comparison to the other elementary schools in Lenoir City in 2014, the population profile of Eaton Elementary School best mirrored the demographic identity of the entire community. With 94.3% White, 0.3% Black, 1.8% Asian or Pacific Islander,

3.6% Hispanic students, 54% of whom were male, this demographic composition is reflective of Loudon County, which is 95.7% white, 1.6% black, and 7.9% Hispanic (United States Census Bureau, 2016). The fact that this school has a population that mirrors that of the county is of benefit, since results of this research can then be considered as in line for the entire county population, not just those attending Eaton Elementary School. Eaton Elementary has a total population of students in grades 1 – 4 of 620; 143 in 1st grade, 155 in 2nd grade, 161 in 3rd grade, and 161 in 4th grade.

Table 2 presents the assessments for academic achievement for Eaton Elementary School and for all the students in the State of Tennessee for the years 2008 through 2012, which includes the time frame for this study (Heyburn, 2014). School systems all across Tennessee use the Discovery Education Assessment (DEA) to monitor students' progress throughout the school year. The student test scores for the DEA fall within four different categories: below basic, basic, proficient, and advanced. The scores are then evaluated on an individual basis to assist students advance if they are below basic or basic. There is also a school compilation report, the Adequate Yearly Progress Score (AYP), that compares each school system with the statewide average as illustrated in table 2. This, as a component of the No Child Left Behind Act serves as a benchmark to keep all school systems in check and functioning at similar abilities and levels, regardless of location or resources.

Table 2

Adequate Yearly Progress Score (AYP) for Eaton Elementary School for years 2008-2012

	Eaton Elementary School				Tennessee Schools Average			
	Percent Below Basic	Percent Basic	Percent Proficient	Percent Advanced	Percent Below Proficient	Percent Basic	Percent Proficient	Percent Advanced
Math								
2012	5.2	35.2	37.4	22.2	16.8	36.0	28.7	18.5
2011	9.0	32.0	38.7	20.6	21.3	37.7	27.0	14.0
2010	-	6.0	44.2	50.2	26.8	38.9	22.7	11.6
2009	3.0	*	35.3	61.8	9.0	*	45.5	45.3
2008	1.0	*	38.0	61.0	9.0	*	45.4	45.6
Reading								
2012	3.3	37.4	43.7	15.6	11.3	38.8	38.1	11.8
2011	4.0	35.0	43.6	17.4	10.8	35.7	39.7	13.8
2010	-	4	37.1	58.7	12.3	37.0	38.3	12.4
2009	0.0	*	35.0	64.7	9.0	*	48.1	42.8
2008	1.0	*	40.0	59.0	8.0	*	46.1	45.9

(Huffman, 2016) *Data not collected for this year for "Basic"

- Data not available or applicable

Each student also completes the Tennessee Comprehensive Assessment Program (TCAP) one time at the end of the school year to evaluate overall student academic achievement. The assessment categories of the DEA coincide with the categories used on the TCAP exam. Below basic means the student who performs at this level has not demonstrated adequate academic ability, thinking capacity, or application of the material and therefore is not ready to progress to the next level of learning. Basic scorers demonstrate a partial mastery of academic content, thinking capacity, and can apply some of what they have learned. These students are only minimally ready to progress to the next level of learning. Proficient is the level educators strive to help students achieve. At this level, students demonstrate a mastery of the academic content and thinking capacity and can apply what they have learned at grade level. The fourth category, advanced, represents students who perform at a superior level of mastery in academic and thinking abilities. They have a great understanding of the application of the information and are well prepared for the next level of learning (McQueen C. , 2015).

Prior to any research at Eaton all teachers from the school were invited to a session that explained the research to be conducted and what role they would play in that research. They were allowed to ask questions and offer suggestions that would aid the process for researcher and teacher alike. The kindergarten teachers thought the research might pose unwelcome challenges to their students and requested not to participate in this study. The wishes of the kindergarten teachers were respected, but all other teachers were willing and agreed to participate.

Study Design

This study utilized an experimental design. When working in educational settings, two types of experimental designs are quite common: the randomized block design and the hierarchical design (Hedges & Rhoads, 2010). The main difference between the two designs exists in how the treatments are randomized. In the randomized block design, individuals within the same cluster (classroom, school, etc.) are assigned different treatments. So with the randomized block design students within a single classroom could possibly have different treatments. The hierarchical design, which this study implements, keeps clusters intact with the treatments. This design is often more widely used, simply because by having each cluster experience with the same treatment, contamination is greatly minimized (Hedges & Rhoads, 2010).

In this study, a meticulous design was created to interpret what impact the stability ball seating truly had on the students. With the stability ball being the independent variable, the dependent variables are math and reading Discovery Test scores and the components of students' behavior through the Teacher Report Form, TRF, of conduct, aggressive and attention deficit hyperactivity disorder syndromes. In order to maximize the results fully, a pre-test and post-test two group experimental design was implemented, utilizing an intervention and control group, as the diagram in Figure 4 illustrates.

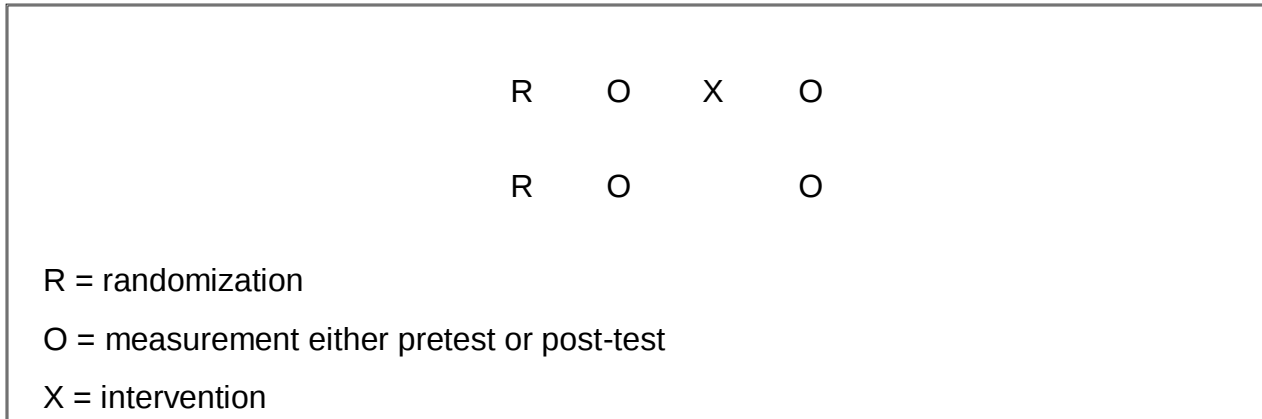


Figure 4: Diagram of pretest-posttest control group experimental design

Through this study design, all students were evaluated on behavior, math, and reading scores prior to implementation of the intervention as well as after implementation. Entire classrooms were randomized to receive the intervention, rather than randomizing students, in order to reduce contamination (e.g., reducing the chance that students not assigned to balls might use the stability balls in the classroom) and to make the study more palatable to teachers. All of the names of teachers were put in a container in grade level groups, and half of the total number of teachers and their classes in each grade were drawn randomly to be in the intervention group. The one student that did not have parental permission for participation was in a classroom drawn to be in the control classroom category. All students in the intervention classrooms were given the option to refuse the seating intervention. No student refused the intervention.

When there was an even number of classes in each grade level, the number of control groups equaled the number of intervention groups. When there was an odd number of classes in a specific grade level, it was randomly decided which group would have one extra class by a coin toss from the school principal.

As shown in Table 1, four classes were assigned to the control group and four to the intervention group among the first-grade classes. Within second grade, there were three control classes and four intervention classes at the beginning of the intervention. Third grade was represented with four classes in the control group and three in the intervention group. Lastly, fourth grade had three control and three intervention classes. All teachers were empowered to opt out should they discover any unforeseen, disruptive issues surface as a result of the study. One teacher in a second-grade class decided to opt out three weeks into the intervention. The data collected for this class were destroyed. Data were collected from 136 students in first grade, 140 students in second grade, 141 in third grade and 80 in fourth grade, for an overall response rate of 80.2%. There were a total of 229 males and 268 females. The complete breakdown is illustrated in table 3. Within each grade, the demographic breakdown was relatively similar.

Table 3
Distribution of students in each grade by control and intervention groups

Grade	Intervention <i>n</i> (%)	Control <i>n</i> (%)	Total <i>n</i> (%)
1	68 (50)	68 (50)	136 (100)
2	85 (61)	55 (39)	140 (100)
3	80 (57)	61 (43)	141 (100)
4	34 (43)	46 (57)	80 (100)
Total	267 (54)	230 (46)	497 (100)

Implementation of the Intervention

Prior to any intervention implementation, IRB approval was sought. Once approval was received in July of 2012 the intervention began in August 2012 with the beginning of the 2012-2013 school year at Eaton Elementary.

As students started their 2012–2013 school year, traditional chairs were placed in all classrooms as in years past in either row seating or group seating. In October - teachers received the Teacher Report Form to report the classroom behavior of each student in each class and were given a month to complete them. They were asked to complete pages 3 and 4 only and to put the student's name and the teacher's name on the front page. Pages 3 and 4 consist of 113 behaviors the teachers were to rank from 0 to 2 with 0 being "Not True (as far as you know)," 1 being "Somewhat or Sometimes True," and 2 being "Very True or Often True." They were all given large manila envelopes and asked to place all completed forms in the envelope and seal it when they were completed. The forms had the names of the children on them so they could be linked to the teacher's classroom and the academic achievement data. Any paperwork with identifiable information was stored in locked filing cabinets in a private office. No access to any such information was made available to anyone at any time.

During the winter break, chairs were removed from intervention classrooms and placed in storage. All desks were then fitted with stability ball seating in sizes teachers specified for students to best fit each one individually. There were three sizes available to accommodate students of varying sizes. As students returned to school, several size adjustments were made, and the teachers continued their classroom management as normal.

As the semester progressed, many questions came via email regarding maintenance and storage of the stability balls. All matters were handled promptly to the satisfaction of all teachers. One teacher, however, decided to abandon the program a few weeks into the intervention. The fact that students could move gradually in a vertical pattern caused the teacher to experience motion sickness, so all chairs were returned and the stability balls were removed. The data that had already been collected for this classroom was not included in any computations.

The assistant principal provided the Discovery Education Assessment scores for reading and math at Eaton for students in grades 1 through 4. The scores for the class dropped from the research project were shredded upon receipt. Data were provided for all three DEA assessments in 2012/2013: in October, January, and April.

Data collection method

At the beginning of each school year, all of Loudon County schools have a registration day when a legal parent or guardian must officially register each student in person. Many forms must be completed for each student each school year to comply with state-mandated school enrollment procedures. Permission was granted by the county and school administrations to include the informed-consent forms for this research for parents or guardians to complete along with the established enrollment forms parent fill out on registration day. Before the in-person registration session, each teacher was trained on the parental-consent form, given information on the purpose of the research again, and reminded how each student may or may not be involved. This took place the week prior to school opening for the academic year. Later that same week, as forms were delivered to each individual teacher, the process for obtaining

informed consent, a review of the research, and the role of the student in this research were reviewed and any remaining questions were answered.

Teachers were trained how to explain the research program to the parents and to indicate that each parent had the option to not consent to the research if they wished, which meant there would be no data collected on their child nor would their child be using the stability ball seating even if that child was in a classroom that was part of the stability ball group. At that time, teachers did not know which classrooms would be intervention classrooms and which classrooms would be control classrooms. Therefore, teachers were not able to convey whether their students would have stability ball seating or not.

The researcher was on the premises all day during the registration hours to answer any questions parents had that teachers were uncertain about answering. No teacher, however, requested clarification or needed help with any parent. As a result of this approach for informed consent, 100% of all forms were returned with only one parent wishing a student be exempt from the program. That parent's wish that her son not be included in the research in any way was upheld and no data was collected for that student, nor did the student use a stability ball.

The researcher also explained how students should utilize the stability balls should their class be selected as an intervention classroom. Specific attention was given to the placement of the stability feet, to be used only when students were not seated on the stability ball. When students were seated, the stability feet were to be on the side of the ball, parallel to the floor to avoid any obstacles while seated.

In October, 2012 teachers were given packets containing enough TRF forms for their entire classroom. The teachers were given a due date to complete the forms and place them all in a large sealed envelope. The school secretary called when a packet had been completed and regular pick-ups were made until all forms were submitted by November 15, 2012. After the teacher turned in the TRF forms, each student was given a code comprising an individually generated alphanumeric identifier. It consisted of the grade of the child, a letter to represent the teacher, and a number for the order in which the teacher submitted each individual form. For example, Jane Doe in Mrs. Jones' class might be 2f10 meaning the 2 is for Jane's grade, second grade, f is the letter randomly assigned to Mrs. Jones, and 10 because Jane's form was the tenth form in the stack as Mrs. Jones completed them. The actual student name was not included in the analysis dataset.

In April, teachers received another set of TRF forms. They were asked to complete these forms keeping the behavior of the students in mind over the past two months, as indicated by the recommendations of the TRF. They once again had one month to complete the surveys following the same procedure as before. All packets were submitted by May 1, 2013. At the end of the study, some teachers requested to keep the stability ball seating in their classrooms, while others allowed students to take the stability ball home.

Measures

Academic achievement was assessed by the reading and mathematics scores of each student on the Discovery Education Assessment (DEA). The Tennessee Comprehensive Assessment Program (TCAP) would seem to be a good choice to

measure academic achievement, since every student completes the same exam at the end of each year. However, the year-end TCAP was not an option since it is only implemented one time a year, usually during the month of April. Since there would only be one score at the end of the year, there would be no baseline established to compare potential academic advancements made while utilizing stability ball seating. Moreover, the TCAP test changes from year to year, measuring different areas with each change. In contrast, the Discovery Education Assessment is given to students periodically throughout the year to measure progress on the key areas of reading and mathematics. It is also consistent from year to year: third-grade students this year are measured on the same criteria that third-grade students had been measured on the previous year. All students take the same assessment three times a year and it measures the same content with each assessment (Amburn, 2012). The test was also already in place, which meant this study imposed less disruption for teachers and classrooms. Consent to utilize these test scores was included in the informed consent process in August.

Teacher Report Form (TRF)

The second set of outcomes is three behavior syndromes measured by the Achenbach System of Empirically Based Assessments (ASEBA), specifically: aggressive behavior, conduct problems, and attention problems. The attention problem syndrome consists of two subcategories, inattention problems and hyperactivity or impulsivity. The term syndromes as used by the ASEBA refers to a similar group of symptoms that consistently occur together (Achenbach & Rescoria, 2001).

ASEBA is a compilation of three assessments commonly used in conjunction with each other to get a view of specific behavioral syndromes from three different

perspectives—the parent, the teacher, and the student. The system gathers data from about the child’s behavior from these three respondents for a more accurate depiction of the student’s behavior syndromes. The Child Behavior Checklist (CBCL) is the component that parents or guardians complete; the Youth Self Report (YSR) is the second assessment, which the student completes; the Teacher Report Form (TRF) is the third assessment, which is completed by the student’s teacher. The reliability and validity of all three instruments have been studied by a number of researchers, and the use of the TRF as an independent measure of student behavior syndromes has been widely validated. (Achenbach & Ruffle, 2000; Bordin, et al., 2012; Gomez, Vance, & Gomez, 2014; Hox, et al., 1999; Ivanova, et al., 2007; So, et al., 2012).

The TRF poses teachers with a series of questions about specific displays of behavior that the student may exhibit in the classroom. The data collected on the TRF generates a profile of syndromes via the scoring software program provided with the instrument. Syndromes include anxious/depressed, withdrawn/depressed, somatic complaints, social problems, thought problems, attention problems, rule-breaking behavior, and aggressive behavior. The behaviors examined for this research are a subset within the externalizing syndromes of rule breaking and aggressive behavior. The complete subset includes affective problems, anxiety problems, somatic problems, attention deficit/hyperactivity problems, oppositional defiant problems and conduct problems. To measure these syndromes, teachers answer 113 questions by indicating “not true as far as you know” (0), “somewhat or sometimes” (1) or “very true or often true” (2) (Achenbach & Rescoria, 2001). Teachers spend approximately 10 minutes to assess each student. The software program provided with the instrument calculates, for

each student, a total score, t score, and percentile for each of the syndromes. For this research, teachers answered all 113 questions but only the scores for items assessing conduct problems, attention problems, and aggressive behavior were used, as these are the behaviors theoretically linked to stability ball seating and most often researched in prior studies (Bill, 2008; Fedewa & Erwin, 2011; Schilling & Schwarz, 2004).

Although the use of a single component of the ASEBA, such as the TRF, is not appropriate for making a clinical diagnoses, it is appropriate to use as a prescreening instrument (Achenbach & Rescoria, 2001). For example, the TRF, used alone, can document conduct problems, but it cannot be used alone to diagnose conduct disorder. The TRF used alone can also validly measure change over time, for example, change in response to an intervention (Achenbach & Rescoria, 2001). The Manual for the ASEBA School-Age Forms and Profiles also recommends that when a single component of the three-part assessment is used, independently computerized scoring should also be used for more accurate results (Achenbach & Rescoria, 2001). An example of the TRF summary report is included in the appendix.

All t scores were generated through a concealed algorithm via the software provided by ASEBA. For this study, data for the three (of the eight) syndromes are used for each student—conduct, attention problems and aggressive behavior-- as those are the behavior characteristics targeted with this intervention.

The following table lists the specific questions that comprise each of the three behavioral syndromes used for this study. Some questions overlap into multiple categories. In total, 38 questions were the focus of this research. Several questions have space to elaborate or specifically describe the behavior described. Within the 38

for this research only one question has a qualitative component. Question 73 asked “Behaves irresponsibly” then asked to “describe” that behavior in more depth. No teacher elaborated on this question.

To prepare teachers for this assessment, a group training was conducted during the week prior to the start of school. Though this is early relative to the actual assessment’s start, this was the only opportunity to address the entire teaching staff and administration at one time. As part of an in-service work day within the school, teachers gathered for an assembly to learn about this research, the role they would play, how to implement the instrument and how to conduct the parental informed consent process.

The parental consent process consisted of a brief explanation of the research. Parents were told that their child may be in a classroom that might be selected for a change of seating. The seating alternative was explained that the stability ball would be provided for each student in the classroom, fitted for the best position of the student for comfort and physical requirements. The seating would be implemented in the second half of the school year, and be in place upon the return from the winter holiday to minimize any classroom disruption. The parent and student could now or at any time ask to have traditional seating replaced and no repercussion or disappointment would result. They were also informed that the teacher could also make that same request on the student’s behalf. Parents were given the opportunity to ask questions of the teacher and were also informed that the researcher was on the premises and would also entertain any questions.

The training in the group setting lasted for approximately two hours with 100%

Table 4

Questions from Teacher Report Form dealing with Conduct Problems, Aggressive Behavior and ADHD Behavior

Conduct Problems		Aggressive Behavior		ADHD Behavior	
Number	Question	Number	Question	Number	Question
16	Cruelty, Bullying, or meanness to others	3	Argues a lot	4	Fails to finish things he/she starts
21	Destroys property belonging to others	6	Defiant, talks back to staff	8	Can't concentrate, can't pay attention for long
26	Doesn't seem to feel guilty after misbehaving	16	Cruelty, bullying or meanness to others	10	Can't sit still, restless, or hyperactive
28	Breaks school rules	19	Demands a lot of attention	15	Fidgets
37	Gets in many fights	20	Destroys his/her own things	22	Difficulty following directions
39	Hangs around with others who get in trouble	21	Destroys property belonging to others	24	Disturbs other pupils
43	Lying or cheating	23	Disobedient at school	41	Impulsive or acts without thinking
57	Physically attacks other people	37	Gets in many fights	53	Talks out of turn
73	Behaves irresponsibly	57	Physically attacks people	67	Disrupts class discipline
82	Steals	68	Screams a lot	78	Inattentive or easily distracted
90	Swearing or obscene language	76	Explosive and unpredictable behavior	93	Talks too much
97	Threatens people	77	Demands must be met immediately, easily frustrated	100	Fails to carry out assigned tasks
101	Truancy or unexplained absence	86	Stubborn, sullen or irritable	104	Unusually loud
		87	Sudden changes in mood or feelings		
		88	Sulks a lot		
		89	Suspicious		
		94	Teases a lot		
		95	Temper tantrums or hot temper		
		97	Threatens people		
		104	Unusually loud		

attendance by all teaching staff and administrators of Eaton Elementary. The staff and administrators for Eaton Elementary consisted of principal and assistant principal, two pre-k teachers and seven teachers in each grade for grades kindergarten through fourth grade. Each section of the instrument was reviewed with specific directions to complete only the Likert-type scale section for each student and provide demographic information, including their name, gender, and birthdate. Teachers were instructed to not fill in any other questions on the TRF. During the training, teachers posed questions primarily about logistics of seating in terms of storage and implementation. No questions were posed regarding the TRF other than due dates.

As the actual assessment date approached, visits to each of the classrooms took place to serve as an opportunity to review the procedure and to also find out if any teachers have developed questions or concerns prior to the beginning of the assessments. As the opportunity for questions with a visit to each classroom took place it was clear each teacher had a strong understanding of the proper protocol with the TRF. Teachers completed the assessments in October and again in May. The October assessment measured the student behavior prior to any intervention as a pretest. The same assessment was then completed again in April and submitted by May 1 after three months of intervention during part of January, February, March and part of April. The teachers were asked to consider the behavior of the student over the past two months, as standard with the TRF (Achenbach T. , 2015).

As with all research instruments, validity and reliability are crucial. The manual for the ASEBA School-Age Forms and Profiles describes a study of 1543 students, both clinically referred and non-referred, with an average age of 11.2 (SD = 3.2). (Achenbach

& Rescoria, Manual for the ASEBA School-Age Forms and Profiles, 2001). The raw TRF scores were regressed using a structural equation model on age, race/ethnicity, clinical referral status and socioeconomic status.

The syndromes focused upon for this study were conduct, aggressive behavior, and attention problems. The variance percentage categories utilized were that of Cohen's 1988 criteria for effect size: small=2-13%; medium=13-26% and large>26% (Achenbach T. , 2015). For those individual syndromes, aggressive behavior had a percent of variance of 19. Attention problems, which also has a set of sub-syndromes of inattention and hyperactivity/impulsivity, had a variance percent of 22. Inattention was 21 and hyperactivity/impulsivity had 14 percent of variance all at a significance of $p<0.01$ showing a medium but significant difference between the referral students with the non-referral students (Achenbach & Rescoria, 2001). With the significance of the variance considered, the percentages of the syndromes utilized for this research were soundly in the medium categorization with the exception of hyperactivity/impulsivity. Hyperactivity/impulsivity still falls within the medium category, but is closer to the small category. The strong p-value verifies the syndrome delineation.

The content validity of the TRF was scrutinized through a study conducted by Verhulst, Koot and Van der Ende in 1994. Their research verified the effectiveness of the TRF in predicting signs of disturbance with the students in their classroom (Verhulst, Koot, & Van der Ende, 1994). With odds ratios of 1.86 for TRF Internalizing Problems, 2.50 for TRF Externalizing and 2.96 for TRF total problems, the Verhulst, et al. study supports the research conducted by Achenbach and Rescoria where all TRF adaptive ability scored significantly lower ($p<.01$) for the referral children than for non-referral

children meaning students with lower adaptive functioning scores display greater challenges with working hard, behaving, learning, and being happy to identify a few (Achenbach T. , 2015).

Convergent validity measures the degree to which two constructs that theoretically should be related are, in fact, related. With Gomez's, et al. research, a multitrait-multimethod (MTMM) approach was used with a correlated trait-correlated method minus 1 [CT-C(M-1)] to measure the convergent validity because the information from the ASEBA comes from multiple sources and measures multiple traits. Archived data from the Academic Child Psychiatry Unit in Australia was used from 294 students of all ages between 5 and 18. Of the total sample, 70.5% were male (Gomez, Vance, & Gomez, 2014). The TRF was compared to the Anxiety Disorders Interview Schedule for Children, Parent version (ADISC-IV-P). Children and parents were interviewed separately for more accurate and honest answers while maintaining discretion (Gomez, Vance, & Gomez, 2014). The reliabilities of all TRF scales were at least moderate at >0.56 (Gomez, Vance, & Gomez, 2014). Results of the CT-C(M-1) model were consistency coefficients of 0.18 to 0.47 for the TRF supporting convergent validity (Gomez, Vance, & Gomez, 2014).

Achenbach and Rescorla, examined the odds ratio of having a diagnosis when exhibiting risk factors. The odds ratio for attention problems was 7.0, meaning those students exhibiting attention problems were seven times more likely to score as having attention problems on the TRF. Both the sub-syndromes of inattention and hyperactivity/impulsivity both received an odds ratio of 5 meaning those students were 5 times more likely to score as such. Lastly aggressive behavior was at 9 meaning those

students were 9 times more likely to score as having aggressive behavior. All of these look at those receiving “deviant” scores on the TRF. Deviant being a term used by Achenbach to categorize the cumulative effect of the syndrome scores toward clinical necessity. To obtain a score of “deviant” the student has a score that is either clinical or borderline clinical according to the TRF scoring mechanism (Achenbach & Rescoria, 2001).

Discovery Education Assessment (DEA)

Test scores for the Discovery Education Assessment were shared for three rounds of testing, October, January and May.

Tests were administered on the same day for each grade level on dates set by the school administration. All students within given grade levels tested simultaneously, but all grades did not test concurrently. Testing was to take place within a specified period of time that the grade level teachers and school administration agreed upon (Amburn, 2012). The Discovery Education Assessment measures academic achievement in reading, mathematics, and science but for the purpose of this research only scores in reading and mathematics were collected.

As a requirement of the No Child Left Behind Act (NCLB) of 2001, the U.S. Department of Education outlined six major criteria for scientifically based research to be used by vendors of educational measurement systems. Those six criteria as listed in NCLB of 2001 are as follows:

- i. “Employs systematic empirical methods that draw on observation and experiment;

- ii. Involves rigorous data analysis that are adequate to test the stated hypotheses and justify the general conclusions drawn;
- iii. Relies on measurements or observational methods that provide reliable and valid data across evaluators and observers, across multiple measurements and observations, and across studies by the same or different investigators;
- iv. Is evaluated using experimental or quasi-experimental designs in which individuals, entities, programs or activities are assigned to different conditions and with appropriate controls to evaluate the effects of the condition of interest, with a preference for random- assignment experiments, or other designs to the extent that those designs contain within-condition or across-condition control.
- v. Ensures experimental studies are presented in sufficient detail and clarity to allow for replication or at minimum offer the opportunity to build systematically on their finding;
- vi. Has been accepted by a peer-reviewed journal or approved by a panel of independent experts through a comparably rigorous, objective and scientific review;" (No Child Left Behind (NCLB) Act of 2001, 2002)

The Discovery Education Assessment documentation states that this assessment meets these criteria and has been approved as an education measure for the state of Tennessee. The content and reporting categories Discovery Education Assessment utilizes are continually updated within the assessment to align with Tennessee state standards. They are as listed:

Tennessee Reading Categories per TCAP

- Content

- Grammar Conventions
- Meaning
- Techniques and Skills
- Vocabulary
- Writing/Organizing
- Writing/Writing Process

Tennessee Mathematics Categories per TCAP

- Number Sense/Number Theory
- Computation
- Algebraic Thinking
- Real World Problem Solving
- Data Analysis and Probability
- Measurement
- Geometry
- Graphs and Graphing

(Shrago & Smith, 2015)

Obtaining research outlining all validity and reliability testing to verify their own claims was not a simple task. However, as a result of a phone interview with Amy Johnson from Discovery Education Assessment, the researcher was provided with a manual of operations which outlined their research for marketing the assessment to school systems (Johnson, 2015). Several research citations were included within this document but when searched, results were not located. A follow up email

correspondence with the Education Commissioner for Tennessee, Candice McQueen, PhD resulted in an attempt on her part to have her staff contact me with the data the office used in approving this assessment for use in Tennessee public schools (McQueen C. P., 2015). To date no other supporting materials have been disclosed.

Within the document provided by Discovery Education Assessment, several reliability and validity study examples were illustrated. The median reading reliability using a Cronbach's alpha was .85, with median sample sizes of 52,628. The median mathematics reliability was .79 with a median sample size of 49,014 (Shrago & Smith, 2015). There was no sample outlined as to the origins of this data, though. They state these data are specific to Tennessee but there was no more mention as to where the data came from exactly, how it was calculated or how it was gathered.

The following tables are described to be the evidence that the questions are consistently measuring reading comprehension and mathematics ability. Test A was given in the fall, Test B in the winter and Test C in the early spring of 2012-2013 school year (Shrago & Smith, 2015). The DEA measures specific content areas throughout the year to allow teachers to cater instruction to maximize the academic achievement for the entire classroom. The same content areas are measured each year in grade/age appropriate instruments. No more information other than the median sample size of 52,628 was disclosed.

A breakdown for Tennessee specific research was provided and illustrated in table 5. This table explains the difference between schools utilizing Discovery Education Assessments and those not implementing this tool. This is the information DEA utilizes in marketing the assessment to the public school system to be used in Tennessee. This

Table 5

Testing reliability coefficients from Cronbach's alpha for reading from Discovery Education Assessment in grades K through 8 for reading and English

2012-2013 Common Core Sample & Reliability Coefficients: Reading						
	Test A		Test B		Test C	
	<i>N</i>	<i>Reliability</i>	<i>N</i>	<i>Reliability</i>	<i>N</i>	<i>Reliability</i>
Kindergarten	44,925	0.71	44,679	0.66	35,555	0.75
Grade 1	83,140	0.70	73,166	0.77	64,948	0.81
Grade 2	77,548	0.84	77,615	0.85	68,531	0.84
Grade 3	55,568	0.87	52,628	0.85	44,207	0.85
Grade 4	59,278	0.85	54,144	0.83	45,864	0.85
Grade 5	59,212	0.87	54,429	0.84	44,982	0.84
Grade 6	58,251	0.88	53,771	0.83	45,156	0.85
Grade 7	56,219	0.84	52,001	0.84	42,548	0.88
Grade 8	55,617	0.86	50,863	0.88	41,699	0.88
English 1	16,159	0.83	11,705	0.86		
English 2	11,587	0.88	7,855	0.85		

(Shrago & Smith, 2015)

Table 6

Testing reliability coefficients from Cronbach's alpha for mathematics from Discovery Education Assessment in grades K through 8 for mathematics and Algebra

2012-2013 Common Core Sample & Reliability Coefficients: Mathematics						
	Test A		Test B		Test C	
	<i>N</i>	<i>Reliability</i>	<i>N</i>	<i>Reliability</i>	<i>N</i>	<i>Reliability</i>
Kindergarten	43,966	0.68	45,279	0.79	24,915	0.82
Grade 1	82,368	0.69	73,745	0.70	64,266	0.79
Grade 2	77,178	0.73	78,542	0.82	67,712	0.83
Grade 3	55,648	0.79	53,792	0.81	44,294	0.83
Grade 4	59,500	0.78	54,446	0.78	45,500	0.81
Grade 5	59,464	0.83	55,300	0.79	45,134	0.83
Grade 6	49,014	0.78	54,176	0.77	44,546	0.81
Grade 7	56,724	0.77	52,205	0.80	42,573	0.80
Grade 8	54,050	0.77	48,688	0.80	38,717	0.80
Algebra 1	11,525	0.72	8,848	0.77		
Algebra 2	12,026	0.71	8,925	0.66		
Geometry	10,529	0.50	7,148	0.63		

(Shrago & Smith, 2015)

study took place during the 2004-2005 school year in the Metro Nashville public schools in Tennessee. It included sixty-five elementary and middle schools with approximately 20,000 students that utilize DEA. There were fifty-two elementary and middle schools with approximately 10,000 students that do not utilize DEA (Shrago & Smith, 2015). This table illustrates the difference between the DEA and non-DEA schools in the improvement of reading and mathematics.

With this table, the percentage point of improvement is greater for those students utilizing the DEA system as opposed to those students not utilizing it. Again, this study is the only statistical information made available to identify any justification the state of Tennessee has to implement this assessment across the state.

Data analysis method

To make sure the groups were a true comparison group, experimental design base level data were compared via a regression through Stata 12. This was calculated to determine that the groups were indeed similar groups with only the difference of the implementation of the intervention. Once that was established the research outcomes could be more clearly linked to the intervention given all else was appropriate. In figure 3 the hierarchical illustration demonstrates how the control group and intervention group randomization resulted. Every effort was made to have balanced numbers within each grade level in each cluster. With odd numbers of classes in three of the four grades that was not always possible.

With a randomized trial with pre and posttests for both the control and intervention groups, selection bias is greatly reduced. This will also aid in the cause and effect relationship identification in this study. Other studies on this subject have yet to

Table 7

Comparison of proficient and/or advanced percentages for DEA and non-DEA schools in Metro Nashville, Tennessee for reading and mathematics in middle and elementary grades.

<i>Grade Level</i>	<i>Status</i>	<i>Year</i>	<i>Reading</i>			<i>Mathematics</i>		
			<i>N</i>	<i>% Prof./ Adv.</i>	<i>% Improve</i>	<i>N</i>	<i>% Prof./ Adv.</i>	<i>% Improve</i>
Combined	DEA	2005	20,190	82.45	10.75	21,549	78.15	7.21
		2004	21,576	71.70		21,738	70.94	
	Non DEA	2005	10,167	88.41	7.45	10,490	83.43	4.36
		2004	10,143	80.96		10,172	79.07	
Elementary	DEA	2005	5,217	86.20	11.71	5,765	80.59	8.33
		2004	5,640	74.49		5,702	72.26	
	Non DEA	2005	5,215	88.40	8.91	5,400	81.81	5.21
		2004	5,309	79.49		5,338	76.60	
Middle	DEA	2005	14,948	81.19	10.42	15,759	77.31	8.77
		2004	15,917	70.77		16,017	70.54	
	Non DEA	2005	4,945	88.41	5.80	5,083	85.19	3.45
		2004	4,831	82.61		4,831	81.74	

(Shrago & Smith, 2015)

Table 8

Comparison of baseline data for gender, age, pretest ADHD, pretest Conduct, pretest Aggressive, January Math DEA and January Reading DEA of control and intervention groups prior to intervention.

	Control Group Mean (s.d.)	Intervention Group Mean (s.d.)	p-value
Male	0.53	0.55	0.72
Age	7.29 (0.99)	7.37 (1.11)	0.46
Pretest ADHD	2.56 (4.50)	4.05 (5.97)	0.002
Pretest Conduct	0.79 (2.76)	0.97 (2.22)	0.45
Pretest Aggressive	1.34 (4.49)	2.15 (5.01)	0.06
Math*	1345.47 (135.49)	1341.91 (142.41)	0.78
Reading*	1391.98 (112.82)	1395.28 (115.30)	0.76

*Math and Reading scores from January

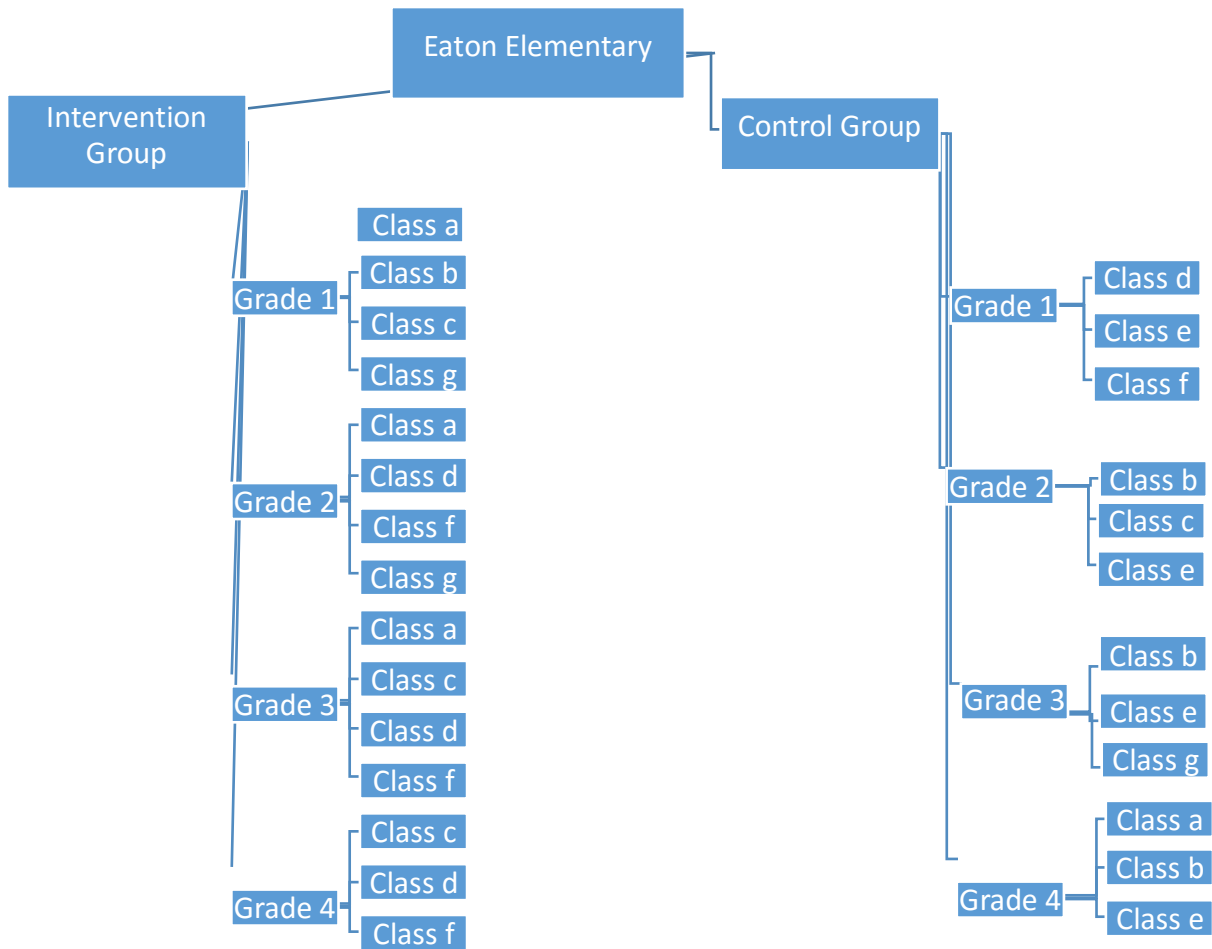


Figure 5: Hierarchical breakdown of study population at Eaton Elementary School in control and intervention groups by grade and class groups

have a solid foundation such as this to base findings. To have results of the analysis of data to be credible, the statistical model must include a means for addressing the hierarchical nature of the data. The multilevel random effects model will accomplish that. It is a straightforward analytic approach for estimating the impact of the treatment for students within a classroom cluster while compensating for the lack of independence in the data for each individual student (Murnane & Willett, 2010). Underestimation of the standard errors is a result of a failure to account for stratification, weighting or clustering of the survey sample (Davern & Strief, 2009)

This study makes two assumptions about the continuous dependent variables being math and reading Discovery Test scores and the components of students' behavior through the Teacher Report Form, TRF, of conduct, aggressive and attention deficit hyperactivity disorder syndromes. First, within each classroom the outcome is normally distributed with a class-specific mean β_{0j} and variance σ_2 . The second assumption is that between classes the outcome is normally distributed with mean γ_{00} and variance τ_{00} .

The level-1 model is as follows:

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij}$$

This represents the within-class level, the student, with β_{0j} = the mean score of classroom j and ε_{ij} = the student level residual.

The level-2 model is then:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}ball_1 + u_{0j}$$

This represents the between-class level, the class as a group, with β_{0j} = the mean score of classroom j , γ_{00} = the mean score of all classrooms, u_{0j} = the difference in the mean scores of all classrooms and classroom j and $\gamma_{01}ball_1$ is the level-2 covariate for the intervention.

The mixed model equation would then be:

$$\gamma_{ij} = \gamma_{00} + \gamma_{01}(ball_1)_{ij} + (u_{0j} + \varepsilon_{ij})$$

Students were clustered according to classroom as opposed to individuals within this study. Since statistical similarities within the cluster usually implies greater similarity for the outcomes than between clusters, a direct impact on the statistical analysis was created on the statistical power. This needs to be taken into consideration when computing data. The intraclass correlation tells us to what extent the students within the cluster are more alike than they are with students between clusters. When a regression was run on the impact of gender, the ICC for gender in relation to math, reading or behavior equaled 0.0002 demonstrating there was no difference with the within clusters or the between clusters in the effect of the ball on reading, math, or behavior as a result of gender.

Stata Statistical Data Analysis software v. 14.1 was used to calculate individual and group level descriptive statistics and regression models. It was also used to calculate the bivariate associations between control and intervention groups prior to the intervention. Four commands were implemented for each variable. To complete the first regression, “xtreg postconduct, i(classid)” is an example of the base data computation

for conduct. This was repeated for aggression, ADHD, reading scores and math scores alike. Following the regression computation, variance was calculated by “* variance sigma e” and then the between and within variances could be calculated, specifically by entering “display e(sigma_u)²” and “display e(sigma_e)²” respectfully. Through these calculations the Model #1 for each variable was complete with the intercept, σ^2_{ϵ} (within variance), σ^2_u (between variance), R^2_{total} (overall variance), and intraclass correlation, ρ .

To interject the intervention, “ball” was added as an independent variable changing the original command to “xtreg postconduct ball, i(classid).” The remaining commands were completed as before recalculating all other output results. The addition of the independent variable “ball” in this computation will give the intercept for the introduction of the intervention on the pretest data for each given dependent variable.

The last cycle of computations added the dependent variable of posttest scores to the equation, “xtreg preconduct ball postconduct, i(classid)”. By adding this final variable, the computations will explain if the changes made in the dependent variables were the result of the intervention or teacher influence.

Power Analysis

Three major components must be taken into consideration to detect the smallest effect that can be detected and shown to be statistically significantly different from zero: the desired significance level of the test, the expected effect size, and the sample size (Hedges & Rhoads, 2010). With multilevel designs such as this research, two additional factors need to be considered, the sample size at each level (e.g., classroom and

students) and the extent of the clustering effects measured by the intraclass correlation coefficient (ICC) (Hedges & Rhoads, 2010).

The choice of significance level for the test depends on the level of type I and type II errors that the investigator is willing to accept. Type I error is when a difference is detected between variables when no difference really exists. When the significance level p , sometimes referred to as α , is 0.05 this signifies that there is a 95% chance of no type I error. The investigator must also select the level of acceptable type II error, detecting no difference when one does actually exist. Type II errors, referred to as a beta error rate or β is generally acceptable at 0.20 or 20%. Thus, $1-\beta = 0.80$. $1-\beta$, also known as the statistical *power*, signifies the chance of a detecting a difference when it actually exists.

The effect size is the difference between the mean of two groups divided by the standard deviation of the outcome from within groups.

$$\delta = \frac{\mu_1 - \mu_2}{\sigma}$$

$\delta = \mu_1 - \mu_2$ is the difference between the mean of group₁ and group₂ and σ is the standard deviation of within group outcomes. When the desired power and sample size are known, it is possible to calculate the effect size that can then be shown to be statistically significantly different from zero. For the multilevel design, you also would need other two pieces of information: sample size at each level and ICC.

With a multilevel design the number of clusters along with the number of students within each cluster must be accounted for as well. So, if an experiment with a control and intervention group has two groups of m clusters with n students in each cluster the sample size becomes $N=2mn$. To compute the operational effect size with no covariate:

$$\Delta^t = \delta \sqrt{\frac{n}{1 + (n - 1)\rho}}$$

Where δ is the effect size, ρ is the ICC and n is the sample size in each cluster.

Note, the operational sample size $N^T=2m$ is smaller than the actual sample size since the clustering level has been added. Using the operational effect size makes it possible to calculate statistical power and sample size requirements for analysis based on clustering by either tables or computer programs. One thing to consider, however, is with this formula an inverse relationship now exists between the ICC (ρ) and the desired effect, *power*. As the ρ increases the power will now decrease. To account for that the formula should now be:

$$\Delta^t = \delta \sqrt{\frac{n}{\frac{1}{n} + \left(n - \frac{1}{n}\right)\rho}}$$

The consideration of the covariates and at which level is their greatest impact is a necessary level in the power calculation that must be considered. So, interjecting the covariates changes the operational sample size to $N^T_A=2m-q$ where q_s is the cluster level covariate and q_w is the individual level covariate. This will reflect the amount of degrees of freedom lost due to the between cluster covariates. Another factor that must not be forgotten is the variances in both between clusters and within clusters. These help dictate the operational effect size and so therefore need to be included in the operational effect size calculation. This now converts the formula to:

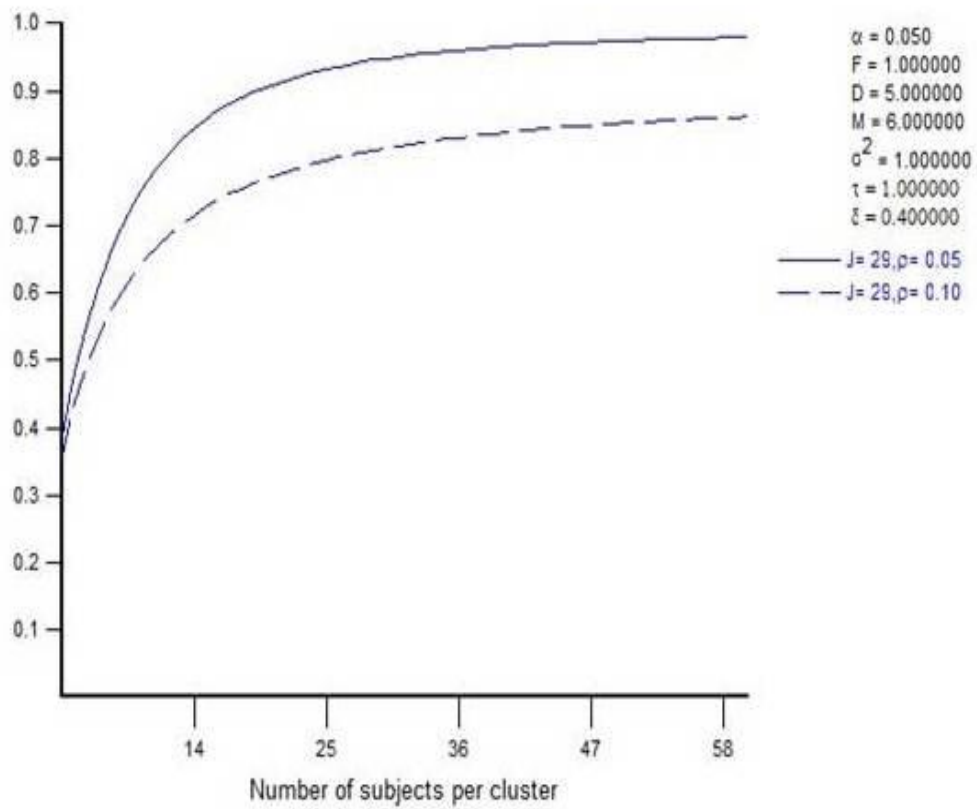
$$\Delta_a^t = \delta \sqrt{\frac{2m}{2m - q_s}} \sqrt{\frac{n}{1 + (n - 1)\rho - [R_w^2 + (nR_s^2 - R_w^2)\rho]}}$$

The first section of the equation is the correction term that is dependent upon the sample size where m is the number of clusters and q is the number of covariates. The second part of the equation is the design effect with the variances of the between and within clusters included.

A sample size that is too small can reduce the reliability of the study and a sample size that is too large may waste time and resources with no added benefit. Getting the correct sample size is just as important as getting the right method or design.

To calculate the power analysis for this study, the William T. Grant Foundation offers the use of the Optimal Design software to compute power for group randomized trials. With this software, the number of clusters or classes entered was 29 and the estimated number of students per class was 25. The alpha was set at 0.05 with the correlation coefficient was estimated at the default of (ρ) at 0.05 as figure 4 illustrates. The student level will be examined, as well, but it is imperative that the power for the group level be strong enough or this will not be relevant for replication.

Here the data compares the impact the stability ball seating had or lack thereof on the math, reading and behavior scores of the students in grades 1-4. The key dependent variables are math and reading Discovery Test scores and the components of students' behavior through the Teacher Report Form, TRF, of conduct, aggressive and attention deficit hyperactivity disorder syndromes. Scores for pretests in all three areas are compared to the posttests with the implementation of the stability ball on the



α = significance level, F = Frequency, D = Duration, M = Occasions,
 σ^2 = class level variance, τ = group level variance, δ = difference in the variances
 J = the number of groups or classes

Figure 6: Power analysis of 29 classes with 25 students in each class.

intervention group. The same score sets for the control group were compared to the intervention group.

The response rate for TRF forms was 98% for the first round of collections with the pretest assessment. Upon follow-up, the remaining two percent of the pre-tests were either turned in or identified as no longer in attendance at Eaton Elementary essentially establishing a 100% response rate for the pretest TRF. The information regarding the math and reading scores from the administration of the DEA were obtained from the principal for the entire student body of Eaton Elementary in grades 1 through 4 so that response rate was 100% as well with a total of 23 classes. The final phase of the TRF assessment with the posttest was more challenging to obtain as teachers were busy with end of school tasks, field trips, and final grading. Teachers asked for extensions in the final submissions resulting in three rounds of submissions for the posttest TRF. The first submission was lacking with only 40% submitted. The second submission brought more with all of the first grade classes and parts of second, third, and fourth. The third round of submissions gave the total of 22 classes roughly 86% of the participants of the pretest total. The drop in student participation has two explanations. First, one class decided to remove the intervention seating in the midst of the research, thus nullifying any pre or post data collected for this group. Second, the number of students transferring from Eaton Elementary school make up the remaining incomplete TRF posttest submissions. This resulted in a total sample size of 623 students in 22 classes. The data for the students that transferred and the students in the class that chose to not complete the study were removed and not included in any calculations.

After the collection of all pre and post TRF forms, pre and post math scores, and pre and post reading scores, data were entered into Excel and imported into the Stata Statistics/Data Analysis software version 12.1.

CHAPTER FOUR

RESULTS

The purpose of this study was to identify any impact of stability ball seating on students in grades 1 through 4 in the general education setting. The key dependent variables are math and reading scores based upon Discovery Test scores and the components of students' behavior assessed through the Teacher Report Form, TRF. The key syndromes from the TRF include conduct, aggressive and attention deficit hyperactivity disorder (ADHD) syndromes. Stata 12 was used for all computations with the exception for the DEA, already computed and submitted by Loudon County School system and the TRF, computed using the software provided by ASEBA. Table 9 outlines baseline data for this intervention.

Table 9
Comparison of pretest and posttest data before and after intervention on ADHD, conduct, aggressive behavior, math and reading scores.

	Control Group Mean (s.d.)	Intervention Group Mean (s.d.)	p-value
Pretest ADHD	2.56 (4.50)	4.05 (5.97)	0.002
Posttest ADHD	1.86 (3.78)	2.58 (4.97)	0.06
Pretest Conduct	0.79 (2.76)	0.97 (2.22)	0.45
Posttest Conduct	0.50 (1.69)	0.78 (2.13)	0.11
Pretest Aggressive	1.34 (4.49)	2.15 (5.01)	0.06
Posttest Aggressive	0.81 (2.81)	1.53 (4.39)	0.03
Math*	1345.47 (135.49)	1341.91 (142.41)	0.78
Math**	1424.44 (230.62)	1427.49 (142.06)	0.87
Reading*	1391.98 (112.82)	1395.28 (115.30)	0.76
Reading**	1450.17 (116.09)	1453.47 (113.79)	0.76

*Math and Reading scores from January

**Math and Reading scores from April

Model #1 in Table 10 is the baseline regression model for conduct posttest. It contains only the intercept. It shows that the mean conduct score for all students is 0.63 (s.e. = 0.10). The within-student variance is 3.58 and the between-class variance is 0.08, resulting in an intraclass correlation of 0.02. This low intraclass correlation coefficient indicates that just a tiny proportion of the variance is due to differences in conduct between classes. In Model #2, the intervention of stability ball seating is introduced. The conduct mean for students without a ball is 0.50 (s.e. = 0.14). The coefficient parameter for the ball is 0.29 with a standard error of 0.20. However, the effect of the ball is not significantly different than zero, meaning the seating alteration has no statistically significant effect on the conduct scores of the students compared to those on traditional seating. The within-student variance remains constant at 3.58 but the between-class variance increases to 0.66. The intraclass correlation remains essentially the same at 0.02. Consistent with the non-significant value of the coefficient for the intervention, the model has a low R^2 , 0.0053.

The pretest value of conduct is introduced as a control variable in Model #3. The coefficient for the stability ball seating is now 0.21, still not significantly different than zero. The coefficient for the pretest is 0.53, indicating that for each one-unit change in the pretest, the posttest increases by 0.53 units. The intraclass correlation in Model #3 is 0.013, which continues to support the finding from Model #1 that the individual student differences account for variability in the conduct scores as opposed to the differences between the groups. With the R^2 at 0.005 with only the ball added into Model #2 it is clear the intervention has very little to no effect for the conduct scores for

students within the intervention group. In Model #3, however the R^2 is 0.49, suggesting that half the variance in conduct score is accounted for by the pretest score.

Table 10
Parameter estimates, standard errors and selected goodness-of-fit statistics for three random effects multilevel models describing the fitted relationship between conduct scores and ball intervention. (n classes = 29)

Fitted Random-Effects Multilevel Model			
	<i>Model #1: The unconditional model</i>	<i>Model #2: Conditional model that contains the main effect of ball</i>	<i>Model #3: Conditional model that adds the main effect of pretests to Model #2</i>
Intercept	.63(.10)***	.50 (.14)***	0.07(0.10)
Ball		.29 (.20)	0.21(0.14)
Pretest			0.53(0.02)***
σ^2_{ϵ}	3.58	3.58	1.84
σ^2_u	0.08	0.66	0.02
R^2_{total}	0.000	0.0053	0.4937
Intraclass correlation, ρ	0.0221	0.0179	0.0134

~ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Aggressive behavior is another outcome variable measured in this study. Table 11 outlines the statistics from the three levels of tests for aggressive behavior; intercept only in Model #1, posttest with the introduction of the stability ball in Model #2 and pretest with stability ball in Model #3. Model #1 shows that the mean aggressive score for all students is 1.13 (s.e. = 0.22). The within-student variance is 12.58 and the between-class variance is 0.53, resulting in the low intraclass correlation of 0.04. This pattern is not much different than the scores in conduct. In Model #2, the intervention of stability ball seating is added to the model. The mean score for aggressive behavior is 0.77 (s.e. = .28) for students without a ball. The parameter of the coefficient for the ball

is 0.78 with a standard error of 0.41. The effect of the ball is 0.03 which shows that there is not any change in aggressive behavior due to the implementation of the stability ball. The within-student variance for the addition of the ball remains constant at 12.58 but the between-class variance decreases to 0.40. The intraclass correlation decreases slightly to 0.03. Consistent with the non-significant value of the coefficient for the intervention, the model has a low R^2 , 0.011.

In Model #3 the pretest score for aggressive behavior is introduced as a control variable. The stability ball seating coefficient is now -0.05. The coefficient for the pretest is 0.61, indicating that for each one-unit increase in the pretest, the posttest increases by 0.61 units. The intraclass correlation in Model #3 is 0.39, which continues to support the finding from Model #1 that the individual student differences account for variability in the aggressive behavior scores as opposed to the differences between the groups. With the R^2 at 0.011 with only the ball added into model #2, it is clear the intervention has very little to no effect for students within the intervention group for aggressive scores. In Model #3, however the R^2 is 0.64, suggesting that the majority of the variance in aggressive score is accounted for by the pretest score.

The last behavior trait measured is attention problems. Table 12 displays Model #1 again as the baseline model that contains only the intercept for attention scores. With attention there is a small difference between the classes as opposed to within the classes as demonstrated by a low intraclass correlation of 0.08. The attention mean score for all students is 2.21 with standard error of 0.31, and the within-student variance and the between-class variance is 17.68 and 1.59 respectively. The intervention of

Table 11

Parameter estimates, standard errors and selected goodness-of-fit statistics for three random effects multilevel models describing the fitted relationship between aggressiveness scores and ball intervention. (n classes = 29)

Fitted Random-Effects Multilevel Model			
	Model #1: The unconditional model	Model #2: Conditional model that contains the main effect of ball	Model #3: Conditional model that adds the main effect of pretests to Model #2
Intercept	1.13(0.22)***	0.77(0.28)~	-0.05(0.18)
Ball		0.78(0.41)**	0.33(0.26)
Pretest			0.61(0.02)***
σ^2_{ϵ}	12.58	12.58	4.51
σ^2_u	.53	.40	0.18
R^2_{total}	0.0000	0.011	0.64
Intraclass correlation, ρ	.04	.031	0.39

~ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

stability ball seating is again introduced in Model #2. The attention mean for students without a ball is 1.85 with a standard error of 0.42. The parameter estimate for the ball is 0.78 with a standard error of 0.62. The between-class variance declines ever so slightly to 1.52 but the within-student variance remains constant at 17.68. The effect of the ball is not significantly different than zero, which shows that utilization of stability ball seating in classroom settings does not impact attention problem behavior in children. The intraclass correlation remains the same at .08. Consistent with the non-significant value of the coefficient for the intervention, the model has a low R^2 , 0.0072.

The control variable of the pretest is again introduced in Model #3. The coefficient for the stability ball seating is now 0.05. The difference in the ball is still not significantly different than zero. The coefficient for the pretest is 0.69, indicating that for each one-unit increase in the pretest, the posttest increases by 0.69 units. Model #3

continues to support the finding that the individual student differences account for variability in attention scores as opposed to the differences between the groups, with an intraclass correlation of 0.09. With the R^2 at 0.007 with only the ball added into Model #2 it is clear the intervention has very little to no effect for the attention scores for students within the intervention group. The variance in attention score is accounted for by the pretest score with the R^2 at 0.69 in Model #3.

Table 12
Parameter estimates, standard errors and selected goodness-of-fit statistics for three random effects multilevel models describing the fitted relationship between attention problems scores and ball intervention. (n classes = 29)

Fitted Random-Effects Multilevel Model			
	<i>Model #1: The unconditional model</i>	<i>Model #2: Conditional model that contains the main effect of ball</i>	<i>Model #3: Conditional model that adds the main effect of pretests to Model #2</i>
Intercept	2.21(0.31)***	1.85(0.42)***	0.05(0.25)
Ball		0.78(0.62)	-0.23(0.36)
Pretest			0.69(0.02)***
σ^2_{ϵ}	17.68	17.68	5.54
σ^2_u	1.59	1.52	0.51
R^2_{total}	0.0000	0.0072	0.69
Intraclass correlation, ρ	0.08	0.08	0.09

~ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Academic achievement is also assessed with this research. The two areas of academics examined are standardized mathematics and reading scores. Model #1 in Table 13 is the baseline model for math that contains only the intercept. The mean math score for all students is 1419.88 with a standard error of 18.87. The within student

variance is 30252.00 and the between class variance is 7480.62, producing in the intraclass correlation of 0.198. This high intraclass correlation demonstrates the impact in student learning is a result of the teacher. In Model #2, the intervention of stability ball seating is introduced. The mean math score for students without a ball is 1420.38 with a standard error of 26.24. The coefficient for the ball is -1.17 with a standard error of 38.62. It appears the stability ball has no impact on the math scores of students, as the effect of the ball is not significantly different than zero. The within-student variance remains constant at 30252.00, and the between-class variance slips to 7863.98. The intraclass correlation increases slightly to .21. Consistent with the non-significant value of the coefficient for the intervention, the model has a low R^2 , .0001.

The pretest math score, from the standardized test taken in October of 2012 is introduced as a control variable in Model #3. The coefficient for the stability ball seating is now 2.56. The difference in the ball is still not significantly different than zero. The coefficient for the pretest is 0.37, indicating that for each one-unit increase in the pretest, the posttest increases by 0.37 units. Though the school administration says they try to avoid ability grouping, there may be some taking place with this group. With the R^2 at 0.0001 with only the ball added into Model #2 it is clear the intervention has very little to no effect for the math scores for students within the intervention group. In Model #3, however the R^2 is 0.116, suggesting the education, not the ball is responsible for any changes taking place.

Model #1 in Table 14 is the baseline model that contains only the intercept for the standardized reading score. It shows that the mean score for all students is 1443.33 (s.e. = 19.97). The within-student variance is 3744.30 and the between-class variance is

10151.21, resulting in the intraclass correlation of 0.73. This higher intraclass correlation coefficient indicates that the majority of the variance is due to differences in reading between classes as opposed to within each classroom. In Model #2, the intervention of

Table 13

Parameter estimates, standard errors and selected goodness-of-fit statistics for three random effects multilevel models describing the fitted relationship between math scores and ball intervention. (n classes = 29)

Fitted Random-Effects Multilevel Model			
	<i>Model #1: The unconditional model</i>	<i>Model #2: Conditional model that contains the main effect of ball</i>	<i>Model #3: Conditional model that adds the main effect of pretest 1 to Model #2</i>
Intercept	1419.88(18.87) ^{***}	1420.38(26.24) ^{***}	920.50(91.50) ^{***}
Ball		-1.17(38.62)	2.56(25.27)
Pretest			0.37(0.07) ^{***}
σ^2_{ϵ}	30252.00	30252.00	29899.52
σ^2_u	7480.62	7863.98	2236.42
R^2_{total}	0.0000	0.0001	0.116
Intraclass correlation, ρ	0.198	0.206	0.070

~ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

stability ball seating is introduced. The mean for reading for students without a ball is 1444.77 with a standard error of 27.78. The coefficient for the ball is -3.12 with a standard error of 41.89. This is not significantly different than zero meaning there is no impact on the reading scores of students from sitting on the stability ball. The within-student variance remains constant at 3744.30 but the between-class variance increases to 10580.32. The intraclass correlation remains the same at 0.74. Consistent with the non-significant value of the coefficient for the intervention, the model has a low R^2 , 0.0003.

The first reading pretest, administered in October 2012, is introduced as a control variable in Model #3. The coefficient for the stability ball seating is now 355.78, still not significantly different than zero. The coefficient for the pretest is 0.99, indicating that for each one-unit increase in the reading pretest, the posttest increases by 0.99 units, an almost perfect correlation. The intraclass correlation in Model #3 at 0.128, which continues to support the finding from Model #1 that the individual student differences account for variability in the reading scores as opposed to the differences between the groups. With the R^2 at 0.0003 with only the ball added into Model #2 it is clear the intervention has very little to no effect for the reading scores for students within the intervention group. In Model #3, however the R^2 is 0.78, suggesting that the majority of the variance in reading score is accounted for by the pretest score.

The second reading pretest, administered in January 2013, is then introduced as a control variable in Model #4. The coefficient for the stability ball seating is now 176.47. The difference in the ball is still not significantly different than zero. The coefficient for the pretest is 0.90, indicating that for each one-unit increase in the pretest, the posttest increases by 0.90 units. The intraclass correlation in Model #3 at 0.008, which, again, continues to support the finding from Model #1 that the individual student differences account for variability in the reading scores as opposed to the differences between the groups. With the R^2 at 0.0003 with only the ball added into model #2 it is clear the intervention has very little to no effect for the reading scores for students within the intervention group. In Model #4, however the R^2 of 0.846, suggests that most of the variance in reading score is accounted for by the pretest score.

Table 14

Parameter estimates, standard errors and selected goodness-of-fit statistics for three random effects multilevel models describing the fitted relationship between reading scores and ball intervention. (n classes = 29)

Fitted Random-Effects Multilevel Model				
	Model #1: The unconditional model	Model #2: Conditional model that contains the main effect of ball	Model #3: Conditional model that adds the main effect of pretest 1 to Model #2	Model #4: Conditional model that adds the main effect of pretest 2 to Model #2
Intercept	1443.33(19.97)***	1444.77(27.78)***	355.78(41.89)***	176.47(27.31)***
Ball		-3.12(40.88)	0.99(8.92)	-0.39(4.56)
Pretest			0.79(0.03)***	0.90(0.02)***
σ^2_{ϵ}	3744.30	3744.30	2295.62	1756.30
σ^2_u	10151.21	10580.32	338.25	14.22
R^2_{total}	0.000	0.0003	0.779	0.846
Intraclass correlation, ρ	0.73	0.739	0.128	0.008

~ $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

CHAPTER FIVE

SUMMARY, CONCLUSION, DISCUSSION, and RECOMMENDATIONS

Chapter five has been organized into four sections. The first section summarizes the entire study. Section two outlines the conclusions drawn from the statistics of the data collection and analysis. The third section discusses the results and implications of this research. Section four concludes this dissertation with recommendations for further research that would be the next logical steps.

Summary

Background

Typically, a classroom is thought to have chairs and desks for students and teachers alike. Well, research has demonstrated that thinking outside the box and offering stability ball seating is beneficial with students often within the CDC classrooms. Some research has even examined to see if the alteration of seating is effective in the workplace. In both cases, studies have seen improvements in productivity with the alternative form of seating (Bill, 2008; Fedewa & Erwin, 2011; Gamache-Hulsmans, 2007; Gregory, Dunk, & Callaghan, 2006; Kingma & van Dieen, 2009; Schilling & Schwartz, Alternative Seating for Young Children with Autism Spectrum Disorder: Effects on Classroom Behavior, 2004). Theories as to why this is the case are hypothesized, however, no significant data analysis has been shared with a definitive reason.

Methodology

This group randomized trial was established to examine the impact stability ball seating had on students in grades 1 through 4 in the general education population on academic achievement in math and reading test scores as measured by the Discovery Education Assessment (DEA). It also examined the impact stability ball seating had on the conduct and aggressive and ADHD behavior of the same students. The behavior variables were measured using the Teacher Report Form (TRF) component of the Achenbach Child Behavior Checklist assessment. This intervention lasted 16 weeks during the second half of the 2011-2012 school year at Eaton Elementary School in Lenoir City, Tennessee.

Parents were presented with a consent agreement at the mandatory registration day. Thanks to the cooperation of the school, 99% response rate. A pretest posttest intervention was utilized using the DEA and TRF as the measures for the pre/post evaluations for academics and behavior variables respectively. Teachers were asked to complete the TRF in October and then again in May. The DEA scores were collected from the regularly scheduled testing dates administered by the school system.

The pre and post tests were evaluated with the control and intervention groups alike to answer the following research questions:

1. What is the impact of stability ball seating, used for 16 weeks in a regular classroom setting, on the conduct, aggressive and attention deficit hyperactivity disorder behavior of students in grades 1 through 4?

2. What is the impact of stability ball seating, used for 16 weeks in a regular classroom setting, on the mathematics scores of students in grades 1 through 4?
3. What is the impact of stability ball, used for 16 weeks in a regular classroom setting, seating on the reading scores of students in grades 1 through 4?

Results

The data analysis revealed that any changes that took place in the behavior of the students in the intervention group could not definitively be attributed to the implementation of the stability ball seating. Also, any improvements in the academics the students achieved during the intervention were not definitively attributed to the stability balls, either. Any scoring variations that occurred between the intervention and control groups were not statistically strong enough to make any claims other than possible teacher intervention, maturity or any number of other variables were responsible.

Conclusion

1. There is no impact of stability ball seating, used for 16 weeks in a regular classroom setting, on the conduct, aggressive and attention deficit hyperactivity disorder behavior of students in grades 1 through 4.
2. There is no impact of stability ball seating, used for 16 weeks in a regular classroom setting, on the mathematics scores of students in grades 1 through 4.
3. There is no impact of stability ball, used for 16 weeks in a regular classroom setting, seating on the reading scores of students in grades 1 through 4.

Discussion

Stability ball seating has been found to have many benefits with student diagnosed with Attention Deficit Hyperactivity Disorder and Autism Spectrum Disorder. The benefits enable a learning to occur that might not have occurred otherwise. This research does not support the cross over intervention into the general education classroom setting.

The reason the study did not find any significant link between seating and outcome is uncertain. Teacher attitude may have an impact. Student excitement for the alternative form of seating may be a distraction, cancelling out any benefit. Maybe an older student population would be more mature, handle the change better and would exhibit greater impact and increased scores. The fact that this did not show definitive impact does not mean this link is still not possible.

Recommendations

This research utilized the student population of one elementary school in Lenoir City, Tennessee during one school year, 2012-2013. The results conclude that there is no reason to believe stability ball seating will impact the behavior or academics of students in grades 1 through 4 of the general education population based on this research. With such positive outcomes with other similar research, a need for future research does exist with the following suggestions:

1. Maintain the study design exhibited in this research, a group randomized trial.

The strength of having pre/post and control/intervention group design adds strength to any findings enabling results to be more definitive.

2. Increase the intervention duration. To ensure any bias, a longer duration with the stability ball seating may be limited. Since this research did not display a negative impact, span the intervention over multiple years for the clearest results.
3. Though the use of specific measure via the TRF help limit teacher bias, an instrument with more rigor, less teacher bias is recommended. Though trained and the risk of bias was explained it is quite difficult to eliminate completely, so every step to minimize is greatly recommended.

REFERENCES

- Achenbach, T. (2015, 2). *ASEBA: Achenbach System of Empirically Based Assessment*. Retrieved from FAQs on Use and Administration of ASEBA Forms: <http://www.aseba.org/support/administration.html>
- Achenbach, T. M., & Rescoria, L. A. (2001). *Manual for the ASEBA School-Age Forms and Profiles*. Burlington, Vermont: Research Center for Children, Youth and Families.
- Achenbach, T. M., & Ruffle, T. M. (2000, August). The Child Behavior Checklist and related forms for assessing behavior/emotional problems and competencies. *Pediatrics in Review*, 21(1), 265-271.
- Al-Eisa, E., Burgadda, S., & Melam, G. R. (2013). Effect of therapy ball seating on learning and sitting discomorts among Saudi female students. *BioMed Research International*, 2013, 1/4. Retrieved from <http://dx.doi.org/10.1155/2013/153165>
- Amburn, M. (2012, May 30). Director of Curriculum. (S. Borders, Interviewer)
- Amstad, H., Bachlin, A., & von Arx, N. (1992). Ball or chair as seating arrangement in the classroom. A prospective study. *Sheweizerische Medizinische Wochenschrift*, 122`21, 811-816.
- Bill, V. N. (2008, May). Effects of Stability Balls on Behavior and Achievement in the Special Education Classroom. *Master of Science in Special Education Action Research Project*. Marshall, Minnesota, United States of America: Southwest Minnesota State University.
- Bordin, I. A., Rocha, M. M., Paula, C. S., Teixeira, M. T., Achenbach, T. M., Rescorla, L. A., & Silveiras, E. F. (2012, October). Child Behavior Checklist (CBCL), Youth Self-Report (YSR), and Teacher's Report Form (TRF): an overview of the development of the original and Brazilian versions. *Revisao Review*, 13-28.
- Budd, H., & Rehling, B. (2014). Simple sensory strategies for future educators. *University of Puget Sound Doctoral Program*, 1-46.
- Capell, B. (2012, October 14). Seating and time on task. *61-683 Research Paper*, 1-47. Maryvill, Missouri, United States of America: Northwest Missouri State University.
- Center for Disease Control and Prevention. (2014, August 3). *NCHS Data Brief*. Retrieved from CDC: Saving lives and prevention 24/7: <http://www.cdc.gov/nchs/data/databriefs/db70.htm>
- Cibulka, J. G. (1987). Theories of education budgeting; Lessons from the management of decline. *Educational Administration Quarterly*, 23(1), 7-40.
- Commerce, U. S. (2014, March). *State and County Quick Facts*. Retrieved March 2014, from United States Census Bureau: <http://www.quickfacts.census.gov/qfd/47/47105.html>
- Cornell, P. (2002). The impact of changes in teaching and learning on furniture and the learning enviornmentet. *New Directions for Teaching and Learning*(92), 33-42.
- Cutler, D. M., & Lleras-Muney, A. (2007, March). Education and health. *National Poverty Center Policy Brief #9*(9), pp. 1-4.

- Daley, D., & Birchwood, J. (2009, October 13). ADHD and academic performance: why does ADHD impact on academic performance and what can be done to support ADHD children in the classroom. *Child: Care, Health, and Development*, 455-464.
- Davern, M., & Strief, J. (2009). *IPUMS User Note: Issues Concerning the Calculation of Standard Errors (i.e. variance estimates) Using IPUMS Data Products*.
- Dunn, R. S., & Dunn, K. J. (1993). Learning styles/teaching styles should they...can they...be matched? *Educational Leadership*, 238224.
- Dunn, R., Beaudry, J. S., & Klavas, A. (2002). Survey of research on learning styles. *California Journal of Science Education*, 2.2, 76-99. California, United States of America.
- DuPaul, G. E. (2006, July 28). Academic interventions for students with ADHD/hyperactivity disorder and a review of literature. *Reading and Writing Quarterly*, 59-82.
- Fedewa, A. L., & Erwin, H. E. (2011, July/August). Stability balls and students with attention and hyperactivity concerns: Implications for on-task and in-seat behavior. *The American Journal of Occupational Therapy*, 65(4), 393-399.
- Fellhoelter, A. (2015, May 12). Second grade instructor. (S. Borders, Interviewer)
- Gamache-Hulsmans, G. (2007). *Stability ball use in the classroom, affect on on-task behavior and handwriting*. Banff, Alberta, Canada: Banff Elementary School.
- Gase, L. N., Kuo, T., Collier, K., Guerrero, L. R., & Wong, M. D. (2014). Assessing the connection between health and education: Identifying potential leverage points for public health to improve student attendance. *American Journal of Public Health*, 104(9), e47-e54. doi:10.2105/AJPH.2014.301977
- Gettinger, M., & Seibert, J. K. (2002). Best practices in increasing academic learning time. *Best Practices in School Psychology*, 4(1), 773-787.
- Gettinger, M., & Seibert, J. K. (2002). Contributions of study skills to academic competence. *School Psychology Review*, 31(3), 350-365.
- Gomez, R., Vance, A., & Gomez, R. M. (2014, May 14). Analysis of the convergent and discriminant validity of the CBCL, TRF, and YSR in clinic-referred sample. *Journal of Abnormal Child Psychology*, 1-32. doi:10.1007/s10802-014-9879-4
- Guild, P. B. (2001, October). Diversity, learning styles and culture. *New Horizons for Learning*, 1-17. Retrieved from <http://education.jhu.edu/PD/newhorizons/strategies/topics/Learning%20Styles/diversity.html>
- Gur, R. C., Sackeim, H. A., & Gur, R. E. (1976). Classroom Seating and Psychopathology: Some Initial Data. *Journal of Abnormal Psychology*, 85(1), 122-124.
- Guradino, C., & Fullerton, C. (2010). Changing behaviors by changing the classroom environment. *Teaching Exceptional Children*, 42(6), 8-13.
- Gutierrez, K. D., & Rogoff, B. (2003, June/July). Cultural ways of learning: individual traits or repertoires of practice. *Educational Researcher*, 32(5), 19-25.

- Harlacher, J. E., Roberts, N. E., & Merrell, K. W. (2006, November/December). Classroom interventions for students with ADHD. *Teaching Exceptional Children*, 39(2), 6-12.
- Harris, M. E., & Tiedemann-Fuller, M. (2010). Frequencies of T-score differences between Achenbach Child Behavior Checklist and Teacher's Report Form summary scales. *Journal of Psychoeducational Assessments*, 28(1), 54-59.
- Hartwig, M. K., Was, C. A., Isaacson, R. M., & Dujnlosky, J. (2012). General knowledge monitoring as a predictor of in-class performance. *British Journal of Educational Psychology*(82), 456-468.
- Hayes, A. F. (2006). A Primer on Multilevel Modeling. *Human Communication Research*(32), 385-410.
- Heyburn, S. E. (2014). *Tennessee Education Program 2.0 Handbook for Computation*. Nashville, TN: Tennessee Department of Education Office of Local Finance.
- Horwitz, R. A. (1979). Psychology of the open classroom. *Review of Educational Research*, 49(1), 71-86.
- Hox, J., Erol, N., Mellenbergh, G. J., Oosterlaan, J., Shalev, R. S., Auerbach, J., . . . Sergeant, J. A. (1999). Syndrome dimensions of the Child Behavior Checklist and the Teacher Report Form: a critical empirical evaluation. *Child Psychology and Psychiatry*, 40(7), 1095-1116.
- Huffman, K. C. (2016, February). *Tennessee Department of Education Score Card*. Retrieved from Tennessee Department of Education:
<http://edu.reportcard.state.tn.us/pls/apex/f?p=200:20:3337482957847404::NO::>
- Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Dumenci, L., Almqvist, F., Bathiche, M., . . . Verhulst, F. C. (2007). Testing the Teachers Report Form syndromes in 20 societies. *School Psychology Review*, 36(3), 468-486.
- Ivory, D. M. (2011, May). *The impact of dynamic furniture on classroom performance; a pilot study*. Retrieved from University of Puget Sound.
- Johnson, A. (2015, February 9). Discovery Education Assessment Representative. (S. Borders, Interviewer)
- Junod, R. E., DuPaul, G. J., Jitendra, A. K., Volpe, R. J., & Cleary, K. S. (2006). Ability to be still and engaged. *Journal of Psychology*(44), 87-104.
- Kern, L., & Clemens, N. H. (2007). Antecedent strategies to promote appropriate classroom behavior. *Psychology in the Schools*, 44(1), 65-75.
- Konold, T. R., Walthall, J. C., & Pianta, R. C. (2004, August). The behavior of child behavior ratings: measurement structure of the Child Behavior Checklist across time, informants and child gender. *Behavioral Disorders*, 29(4), 372-383.
- Kunst, A. E., & Mackenbach, J. P. (1994, June). The size of mortality differences associated with educational level in nine industrialized countries. *American Journal of Public Health*, 84(6), 932-937.

- Legewie, J., & DiPrete, T. A. (2012, April 4). School context and the gender gap in educational achievement. *American Sociological Review*, 77(3), 463-485.
- Leung, P. W., Kwong, S. L., Tang, C. P., Ho, T. P., Hung, S. F., Lee, C. C., . . . Lieu, W. S. (2006, September). Test re-test reliability and criterion validity of Chinese version of CBCL, TRF and YSR. *Journal of child Psychology and Psychiatry*, 47(9), 970-973.
- Liljequist, L. (2002). Strategies used by teachers to rate student behavior. *Psychology in the Schools*, 39(1), 77-84.
- Lopes, P. N., Mestre, J. M., Guil, R., Kremenitzer, J. P., & Salovey, P. (2012, August). The role of knowledge and skills for managing emotions in adaptation to school: Social behavior and misconduct in the classroom. *American Educational Research Journal*, 49(4), 710-742.
- Lynch, S. M. (2003, May). Cohort and life-course patterns in the relationship between education and health: a hierarchical approach. *Demography*, 40(2).
- Marx, A., Fuhrer, U., & Hartig, T. (2000). Effects of classroom seating arrangements on children's question-asking. *Learning Environment Research*, 2, 249-263.
- Marx, A., Fuhrer, U., & Hartig, T. (2000). Effects of classroom seating arrangements on children's question-asking. *Learning Environments Research*(2), 249-263.
- McConaughy, S. H. (1993). Advances in empirically based assessment of children's behavioral and emotional problems. *School Psychology Review*, 22(2), 285-308.
- McQueen, C. (2015, 2 16). *Assessment*. Retrieved from Tennessee Department of Education: <http://www.state.tn.us/education/assessment/index.shtml>
- McQueen, C. P. (2015, February 23). Commissioner, Tennessee Department of Education. (S. Borders, Interviewer) Nashville, TN.
- Medina, J. (2009). *Brain Rules*. Seattle, WA: Pear Press.
- Merritt, J. M. (2014, January). Alternative seating for young children; Effects on learning. *American International Journal of Contemporary Research*, 4(1), 12-18.
- Montello, D. R. (1988). Classroom seating location and its effect on course achievement, participation, and attitudes. *Journal of Environmental Psychology*(8), 149-157.
- Murnane, R. J., & Willett, J. B. (2010). *Methods Matter: Improving Causal Inference in Educational and Social Science Research*. Oxford Press.
- No Child Left Behind (NCLB) Act of 2001. (2002). *Pub. L., No. 107-110*(115).
- O'Neill, K. B., & Liljequist, L. (2002). Strategies used by teachers to rate student behavior. *Psychology in the Schools*, 39(1), 77-85.
- Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008, December). Learning styles: concepts and evidence. *Journal of the Association for Psychological Science*, 9(3), 105-119.

- Patton, J. E., Snell, J., Knight, W. J., & Gerken, K. (2001). *A Survey Study of Elementary Classroom Seating Designs*. Washington, D.C.: Annual Convention of the National Association of School Psychologists. Retrieved March 2014
- Presley, C. (2014, August 10). Business Manager for Loudon County Department of Education. (S. Borders, Interviewer)
- Research Center for Children, Y. a. (2014). The Worldwide Standard for Evidence Based Assessment. *Winter 2014 Catalog*. Burlington , Vermont, United States of America.
- Roussos, A., Karantanos, C., Richardson, C., Hartman, C., Karajiannis, D., Kyprianos, S., . . . Zoubou, V. (1999). Achenbach's Child Behavior Checklist and Teachers' Report Form in a normative sample of Greek children 6-12 years old. *European Child & Adolescent Psychiatry*(8), 165-172.
- Schilling, D. L., & Schwartz, I. S. (2004, August). Alternative Seating for Young Children with Autism Spectrum Disorder: Effects on Classroom Behavior. *Journal of Autism and Development Disorders*, 34(4), 423-432.
- Schilling, D. L., Washington, K., Billingsley, F. F., & Deitz, J. (2003, September/October). Classroom seating for children with attention deficit hyperactivity disorder: Therapy balls versus chairs. *The American Journal of Occupational Therapy*, 57(5), 534-541.
- Shrago, J., & Smith, M. (2015). Discovery Education Assessment Research. Hershey, Pennsylvania: DiscoveryEducation.com.
- Sines, J. S. (2009). The perceptions of children following participation in a yoga and mindful program: a qualitative study. *Doctoral and Masters' Thesis Program*, 1-62.
- So, P., Greaves-Lord, K., van der Ende, J., Verhulst, F. C., Rescorla, L., & de Nijs, P. F. (2012). Using the Child Behavior Checklist and the Teacher Report Form for identification of children with autism. 17(5), pp. 595-607.
- Spaulding, S. A., Irvin, L. K., Horner, R. H., May, S. L., Emeldi, M. ., Tobin, T. J., & Sugai, G. (2010, April). Schoolwide social-behavioral climate, student problem behavior and related administrative decisions. *Journal of Positive Behavior Interventions*, 12(2), 69-85.
- Vaughn, M. L., Riccio, C. A., Hynd, G. W., & Hall, J. (1997, October 18). Diagnosing ADHD (predominantly inattentive and combine type subtypes): Discriminant validity of the behavior assessment system for children and the Achenbach parent and teacher rating scale. *Journal of clinical Child Psychology*, 26(4), 349-357.
- Wannarka, R., & Ruhl, K. (2008). Seating arrangements that promote positive academic and behavioural outcomes; a review of empirical research. *Support for Learning*, 23(2), 89-93.
- Wasnock, D. P. (2010). *Classroom environment: emphasis on seating arrangement*. St. John Fisher College, Mathematical and Computer Sciences. Fisher Digital Publications. Retrieved from http://fisherpub.sjfc.edu/mathcs_etd_masters
- Weinstein, C. S. (1979). The physical environment of the school: a review of the research. *Review of Educational Research*, 49(4), 577-610.

- Weinstein, C. S. (1992). *Designing the instructional environment; Focus on seating*. Proceedings of selected research and development presentation at the convention of the association for educational communications and technology.
- Whalon, K. J., & Hart, J. E. (2010, February 26). Children with ASD and literacy instruction; an exploratory study of elementary inclusion settings. *Remedial and Special Education*, 1-13.
- Wu, W.-L., Want, C.-C., Chen, C.-H., Lai, C.-L., Yang, P.-C., & Guo, L.-Y. (2012). Influence of therapy ball seat on attentional ability in children with attention deficit/hyperactivity disorder. *Journal of Physical Therapy Science*, 24, 1177-1182.
- Youngstrom, E. A., Joseph, M. F., & Greene, J. (2008). Comparing the psychometric properties of multiple teacher report instruments as predictors of bipolar disorder in children and adolescents. *Journal of Clinical Psychology*, 64(4), 382-401.

APPENDICES

APPENDIX 1

TRF/6-18 - Narrative Report & Critical Items

Scored using T scores for ASEBA Standard

ID: 2e17-1

Birth Date: / /

Informant:

Name:

Gender: Female

Clinician:

Age: 7

Date Filled: 02/01/2012

The Teacher Report Form (TRF) was completed by _____ classroom teacher to obtain her perceptions of adaptive functioning and problems.

_____ classroom teacher rated _____'s performance in one subject at somewhat below grade level, and five subjects at grade level. _____'s classroom teacher rated _____'s working hard about average, behaving appropriately about average, learning about average, and happy about average compared to typical students of the same age.

_____ Academic Performance score was in the normal range for teachers' ratings on girls aged 6 to 11. _____ Total Adaptive Functioning score was in the normal range.

On the TRF problem scales, _____ Total Problems, Internalizing, and Externalizing scores were all in the normal range for girls aged 6 to 11. Scores on all rated syndrome scales were in the normal range. On the Attention Problems subscales, _____ scores for both Inattention and Hyperactivity-Impulsivity were in the normal range.

On the DSM-oriented scales, _____ scores on all rated scales were in the normal range. On the Attention Deficit/Hyperactivity subscales, _____ scores for both Inattention and Hyperactivity-Impulsivity were in the normal range.

On the 2007 scales, _____ scores on all rated scales were in the normal range.

Critical Items

In addition to the scale scores, it is important to consider scores on individual problem items. Because they may raise particular challenges for management, it is especially important to note the problems listed below that were reported with scores of 1 or 2. Look at comments made by the informant on the form in relation to these problems to obtain more information about risks associated with the problems and the contexts in which the problems occur.

Score	Problem Item	Score	Problem Item
0	18. HarmSelf	0	70. SeesThings
0	40. HearsThings	0	91. TalkSoleldo
0	57. Attacks	0	105. AtcDrugs

TRE/6-18 - Syndrome Scale Scores for Girls 6-11 Scored Using T Scores for ASEBA Standard

Page 2 of 5

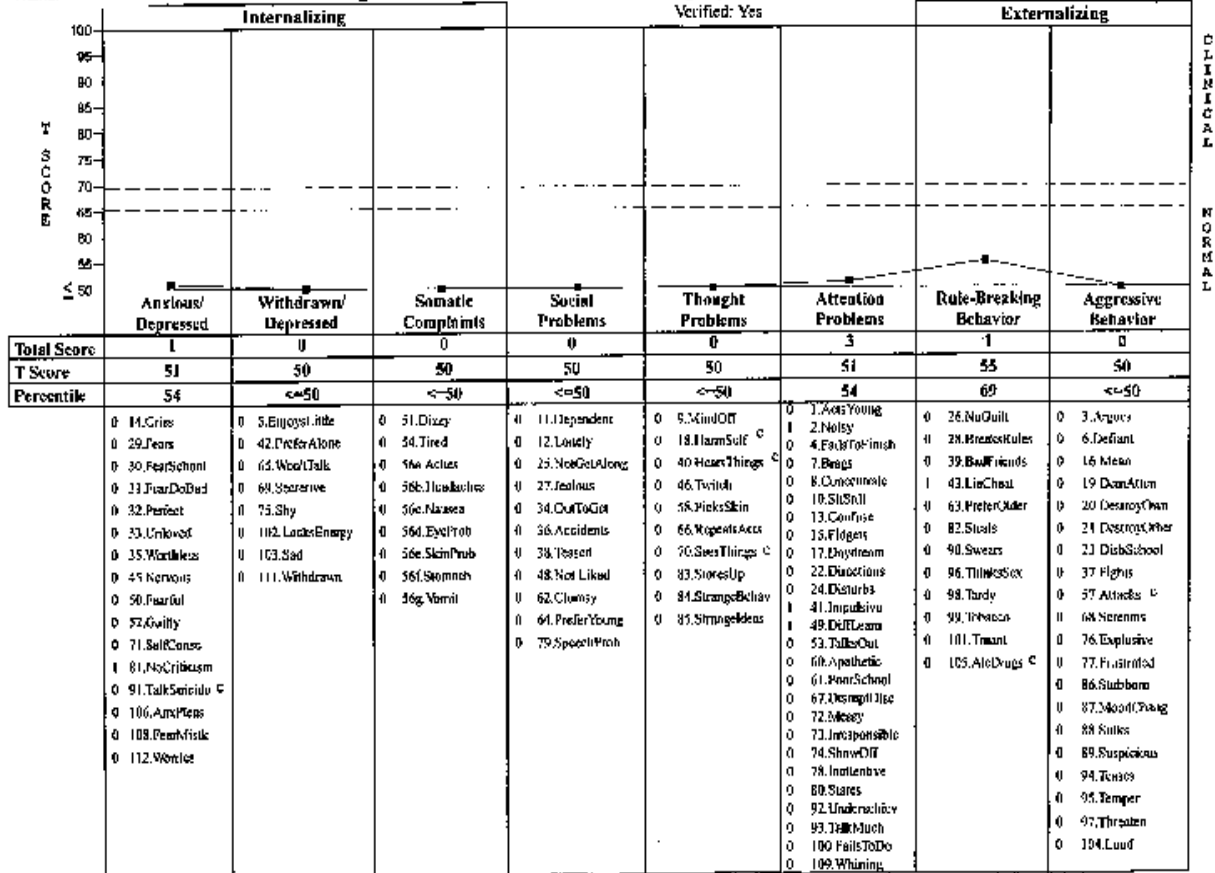
ID: 2e17-f
Name: *

Gender: Female
Age: 7

Date Filled: 02/01/2012
Birth Date: / /

Clinician:
Agency:
Verified: Yes

Informant: I
Relationship: Classroom Teacher (T)



B = Borderline clinical range; C = Clinical range

Broken lines = Borderline clinical range

TRF/6-18 - Internalizing, Externalizing, Total Problems, Other Problems for Girls 6-11
 Scored using T scores for ASEBA Standard

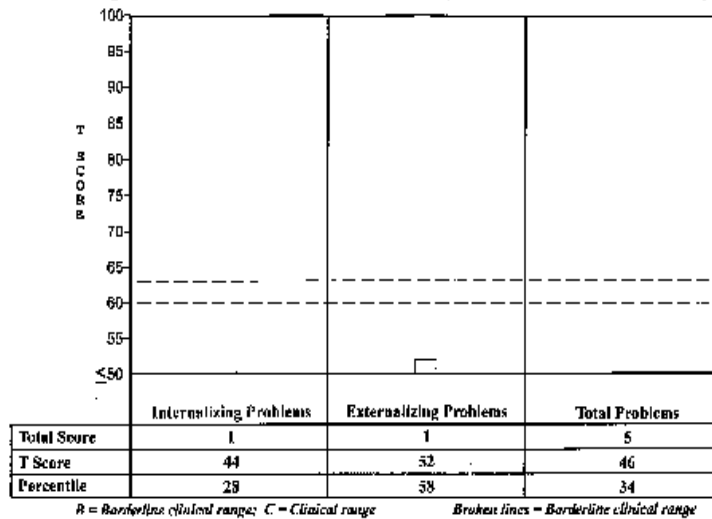
Page 3 of 5

ID: 2e17-1
 Name: _____

Gender: Female Date Filled: 02/01/2012 Clinician:
 Age: 7 Birth Date: / / Agency:

Informant:
 Relationship: Classroom Teacher (F)

Attention Problems Subscale: Inattention (I) Hyperactivity-Impulsivity (H-I)		
I	Attention Problems	H-I
0	1. Acts Young	0
0	2. Noisy	0
0	4. Fails To Finish	0
0	7. Drags	0
0	8. Confronts	0
0	10. Not Still	0
0	13. Confused	0
0	15. Fidgets	0
0	17. Daydreams	0
0	22. Disoriented	0
0	24. Disruptive	0
0	41. Impulsive	0
0	48. Diff. Learn	0
0	53. Tires Out	0
0	60. Appear. Go	0
0	64. Poor School	0
0	67. Disrupt. Dis.	0
0	72. Messy	0
0	73. Inexp. Able	0
0	74. Shows Off	0
0	78. Inattentive	0
0	80. Stares	0
0	92. Uncon. Clav	0
0	93. Talk Much	0
0	100. Fails To Do	0
0	109. Whining	0
1	Total Score	2
≤ 50	Percentile	69
Borderline = 93rd-97th %ile Clinical = >97th %ile		



Other Problems	
0	44. Bili. Nail
0	47. Overconfident
0	55. Overweight
0	56h. Other Phys
0	59. Sleeps Class
0	107. Dislike Self
0	110. Unclean
0	113. Other Prob

Copyright 2007 T.M. Achenbach
 ASEBA, University of Vermont
 1 South Prospect St.
 Burlington, VT 05401-3456
 www.ASEBA.org

ADM Version 9

TRF/6-18 - DSM-Oriented Scales for Girls 6-11 Scored Using T Scores for ASEBA Standard

Page 4 of 5

ID: 2e17-1
Name:

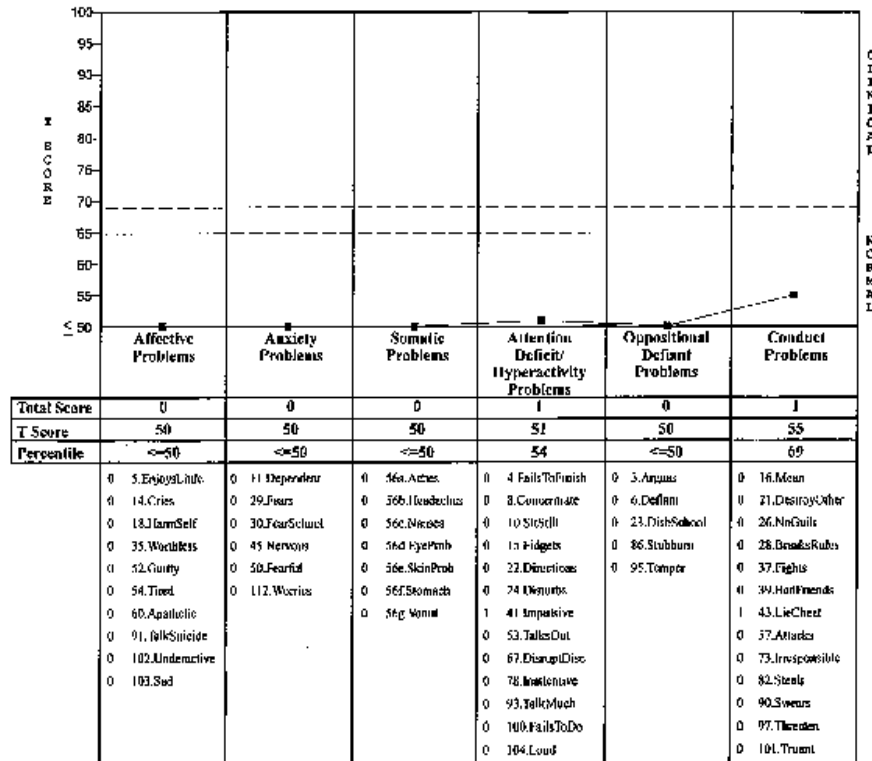
Gender: Female
Age: 7

Date Filled: 02/01/2012
Birth Date: / /

Clinician:
Agency:

Informant: I

Relationship: Classroom Teacher [F]



B = Borderline clinical range; C = Clinical range

Broken lines = Borderline clinical range

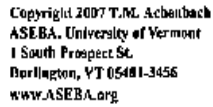
Attention Deficit/Hyperactivity Problems Subscales:		
Inattention (I)		
Hyperactivity-Impulsivity (H-I)		
I	AD/H Problems	H-I
0	4.FailsToFinish	0
0	8.Concentrate	0
0	10.Still	0
0	15.Fidgets	0
0	22.Distractious	0
0	24.Disturbs	0
0	41.Impulsive	1
0	53.TalksOut	0
0	67.DisruptDisc	0
0	78.Inattentive	0
0	81.TalkMuch	0
0	100.FailsToDo	0
0	104.Loud	0
0	Total Score	1
≤ 50	Percentile	63

Borderline = 93rd-97th %ile
Clinical = >97th %ile

Copyright 2007 T. M. Achenbach

Page 5 of 5

Informant:
Relationship: Classroom Teacher {F}



APPENDIX 2


2801

TEACHER'S REPORT FORM FOR AGES 6-18

For office use only
ID #

Your answers will be used to compare the pupil with other pupils whose teachers have completed similar forms. The information from this form will also be used for comparison with other information about this pupil. Please answer as well as you can, even if you lack full information. Scores on individual items will be combined to identify general patterns of behavior. Feel free to print additional comments beside each item and in the spaces provided on page 2. **Please print, and answer all items.**

PUPIL'S FULL NAME	First _____ Middle _____ Last _____	PARENTS' USUAL TYPE OF WORK, even if not working now (Please be specific — for example, auto mechanic, high school teacher, homemaker, laborer, lathe operator, shoe salesman, army sergeant.)
PUPIL'S GENDER ☐ Boy ☐ Girl	Teacher Name _____ Student Name _____	FATHER'S TYPE OF WORK _____ MOTHER'S TYPE OF WORK _____
TODAY'S DATE Mo _____	Day _____ (or F _____)	THIS FORM FILLED OUT BY: (print your full name) _____
GRADE IN SCHOOL _____	NAME AND ADDRESS OF SCHOOL _____ _____ _____	Your gender: ☐ Male ☐ Female Your role at the school: ☐ Classroom Teacher ☐ Counselor ☐ Special Educator ☐ Administrator ☐ Teacher's Aide ☐ Other (specify): _____

I. For how many months have you known this pupil? _____ months

II. How well do you know him/her? 1. ☐ Not Well 2. ☐ Moderately Well 3. ☐ Very Well

III. How much time does he/she spend in your class or service per week? _____

IV. What kind of class or service is it? (Please be specific, e.g., regular 5th grade, 7th grade math, learning disability, counseling, etc.) _____

V. Has he/she ever been referred for special class placement, services, or tutoring?
☐ Don't Know 0. ☐ No 1. ☐ Yes — what kind and when? _____

VI. Has he/she repeated any grades? ☐ Don't Know 0. ☐ No 1. ☐ Yes — grades and reasons: _____

VII. Current academic performance — list academic subjects and check box that indicates pupil's performance for each subject:

Academic subject	1. Far below grade	2. Somewhat below grade	3. At grade level	4. Somewhat above grade	5. Far above grade
1. _____	☐	☐	☐	☐	☐
2. _____	☐	☐	☐	☐	☐
3. _____	☐	☐	☐	☐	☐
4. _____	☐	☐	☐	☐	☐
5. _____	☐	☐	☐	☐	☐
6. _____	☐	☐	☐	☐	☐

Be sure you answered all items. Then see other side.

Copyright 2001 T. Achienbach
 ASEBA, University of Vermont
 1 South Prospect St., Burlington, VT 05401-3456
 www.ASEBA.org

UNAUTHORIZED COPYING IS ILLEGAL


6-1-01 Edition - 301

Please print. Be sure to answer all items.

VIII. Compared to typical pupils of the same age:	1. Much less	2. Somewhat less	3. Slightly less	4. About average	5. Slightly more	6. Somewhat more	7. Much more
1. How hard is he/she working?	☐	☐	☐	☐	☐	☐	☐
2. How appropriately is he/she behaving?	☐	☐	☐	☐	☐	☐	☐
3. How much is he/she learning?	☐	☐	☐	☐	☐	☐	☐
4. How happy is he/she?	☐	☐	☐	☐	☐	☐	☐

IX. Most recent achievement test scores (optional):

Name of test	Subject	Date	Percentile or grade level obtained

X. IQ, readiness, or aptitude tests (optional):

Name of test	Date	IQ or equivalent scores

Does this pupil have any illness or disability (either physical or mental)? ☐ No ☐ Yes— please describe:

What concerns you most about this pupil?

Please describe the best things about this pupil:

Please feel free to write any comments about this pupil's work, behavior, or potential, using extra pages if necessary.

Please print. Be sure to answer all items.

Below is a list of items that describe pupils. For each item that describes the pupil *now or within the past 2 months*, please circle the 2 if the item is **very true or often true** of the pupil. Circle the 1 if the item is **somewhat or sometimes true** of the pupil. If the item is **not true** of the pupil, circle the 0. Please answer all items as well as you can, even if some do not seem to apply to this pupil.

0 = Not True (as far as you know)

1 = Somewhat or Sometimes True

2 = Very True or Often True

0 1 2	1. Acts too young for his/her age	0 1 2	34. Feels others are out to get him/her
0 1 2	2. Hums or makes other odd noises in class	0 1 2	35. Feels worthless or inferior
0 1 2	3. Argues a lot	0 1 2	36. Gets hurt a lot, accident-prone
0 1 2	4. Fails to finish things he/she starts	0 1 2	37. Gets in many fights
0 1 2	5. There is very little that he/she enjoys	0 1 2	38. Gets teased a lot
0 1 2	6. Defiant, talks back to staff	0 1 2	39. Hangs around with others who get in trouble
0 1 2	7. Bragging, boasting	0 1 2	40. Hears sounds or voices that aren't there (describe): _____
0 1 2	8. Can't concentrate, can't pay attention for long	0 1 2	41. Impulsive or acts without thinking
0 1 2	9. Can't get his/her mind off certain thoughts; obsessions (describe): _____	0 1 2	42. Would rather be alone than with others
0 1 2	10. Can't sit still, restless, or hyperactive	0 1 2	43. Lying or cheating
0 1 2	11. Clings to adults or too dependent	0 1 2	44. Bites fingernails
0 1 2	12. Complains of loneliness	0 1 2	45. Nervous, high-strung, or tense
0 1 2	13. Confused or seems to be in a fog	0 1 2	46. Nervous movements or twitching (describe): _____
0 1 2	14. Cries a lot	0 1 2	47. Overconforms to rules
0 1 2	15. Fidgets	0 1 2	48. Not liked by other pupils
0 1 2	16. Cruelty, bullying, or meanness to others	0 1 2	49. Has difficulty learning
0 1 2	17. Daydreams or gets lost in his/her thoughts	0 1 2	50. Too fearful or anxious
0 1 2	18. Deliberately harms self or attempts suicide	0 1 2	51. Feels dizzy or lightheaded
0 1 2	19. Demands a lot of attention	0 1 2	52. Feels too guilty
0 1 2	20. Destroys his/her own things	0 1 2	53. Talks out of turn
0 1 2	21. Destroys property belonging to others	0 1 2	54. Overtired without good reason
0 1 2	22. Difficulty following directions	0 1 2	55. Overweight
0 1 2	23. Disobedient at school	0 1 2	56. Physical problems without known medical cause:
0 1 2	24. Disturbs other pupils	0 1 2	a. Aches or pains (not stomach or headaches)
0 1 2	25. Doesn't get along with other pupils	0 1 2	b. Headaches
0 1 2	26. Doesn't seem to feel guilty after misbehaving	0 1 2	c. Nausea, feels sick
0 1 2	27. Easily jealous	0 1 2	d. Eye problems (not if corrected by glasses) (describe): _____
0 1 2	28. Breaks school rules	0 1 2	e. Rashes or other skin problems
0 1 2	29. Fears certain animals, situations, or places other than school (describe): _____	0 1 2	f. Stomachaches
0 1 2	30. Fears going to school	0 1 2	g. Vomiting, throwing up
0 1 2	31. Fears he/she might think or do something bad	0 1 2	h. Other (describe): _____
0 1 2	32. Feels he/she has to be perfect		
0 1 2	33. Feels or complains that no one loves him/her		

PAGE 3 Be sure you answered all items. Then see other side.

Please print. Be sure to answer all items.

0 = Not True (as far as you know)

1 = Somewhat or Sometimes True

2 = Very True or Often True

- 0 1 2 57. Physically attacks people
0 1 2 58. Picks nose, skin, or other parts of body (describe): _____

- 0 1 2 59. Sleeps in class
0 1 2 60. Apathetic or unmotivated
0 1 2 61. Poor school work
0 1 2 62. Poorly coordinated or clumsy
0 1 2 63. Prefers being with older children or youths
0 1 2 64. Prefers being with younger children
0 1 2 65. Refuses to talk
0 1 2 66. Repeats certain acts over and over; compulsions (describe): _____

- 0 1 2 67. Disrupts class discipline
0 1 2 68. Screams a lot
0 1 2 69. Secretive, keeps things to self
0 1 2 70. Sees things that aren't there (describe): _____

- 0 1 2 71. Self-conscious or easily embarrassed
0 1 2 72. Messy work
0 1 2 73. Behaves irresponsibly (describe): _____

- 0 1 2 74. Showing off or clowning
0 1 2 75. Too shy or timid
0 1 2 76. Explosive and unpredictable behavior
0 1 2 77. Demands must be met immediately, easily frustrated
0 1 2 78. Inattentive or easily distracted
0 1 2 79. Speech problem (describe): _____

- 0 1 2 80. Stares blankly
0 1 2 81. Feels hurt when criticized
0 1 2 82. Steals
0 1 2 83. Stores up too many things he/she doesn't need (describe): _____

- 0 1 2 84. Strange behavior (describe): _____

- 0 1 2 85. Strange ideas (describe): _____

- 0 1 2 86. Stubborn, sullen, or irritable
0 1 2 87. Sudden changes in mood or feelings

- 0 1 2 88. Sulks a lot
0 1 2 89. Suspicious

- 0 1 2 90. Swearing or obscene language
0 1 2 91. Talks about killing self

- 0 1 2 92. Underachieving, not working up to potential
0 1 2 93. Talks too much

- 0 1 2 94. Teases a lot
0 1 2 95. Temper tantrums or hot temper

- 0 1 2 96. Seems preoccupied with sex
0 1 2 97. Threatens people

- 0 1 2 98. Tardy to school or class
0 1 2 99. Smokes, chews, or sniffs tobacco

- 0 1 2 100. Fails to carry out assigned tasks
0 1 2 101. Truancy or unexplained absence

- 0 1 2 102. Underactive, slow moving, or lacks energy
0 1 2 103. Unhappy, sad, or depressed

- 0 1 2 104. Unusually loud
0 1 2 105. Uses alcohol or drugs for nonmedical purposes (*don't* include tobacco) (describe): _____

- 0 1 2 106. Overly anxious to please
0 1 2 107. Dislikes school

- 0 1 2 108. Is afraid of making mistakes
0 1 2 109. Whining

- 0 1 2 110. Unclean personal appearance
0 1 2 111. Withdrawn, doesn't get involved with others

- 0 1 2 112. Worries
0 1 2 113. Please write in any problems the pupil has that were not listed above.

- 0 1 2 _____
0 1 2 _____
0 1 2 _____

VITA

Sheila R. Borders has worked in the field of public health for a number of years during her tenure with the University of Tennessee Extension Service. Ms. Borders began her work with UT Extension in 1995 working with you in Monroe County Tennessee. During this season of employment, she was responsible for completing an annual needs assessment, deliver programs according to identified needs, and evaluate the impact of the programs in an effort to improve the lives and wellbeing of the youth within her program. While working in Monroe County she developed partnerships, established an environmental health initiative with the creation of a field trip classroom and curriculum for all second through fourth graders to learn and understand the impact the environment plays on personal health and wellness. Ms. Borders completed her Master's Degree while employed in Monroe County in Human Resource Development.

In 2001 Ms. Borders moved to Loudon County where her program with UT Extension changed from a primarily youth audience to an adult audience. Through this move, she was able to conduct more programming on adult health programs throughout the county based on the needs of the local residents. As a result of continued needs assessment, programs on chronic conditions such as arthritis, fibromyalgia and diabetes are conducted annually. Another program Ms. Borders attained specialized training to conduct is Tai Chi for Health. Through this program, Ms. Borders works with people of all ages motivated by controlling arthritis pain, loss of balance, or stress and are able to find relief through the slow controlled movements of this exercise.

Currently as a candidate in the doctoral program with the Department of Public Health at the University of Tennessee, Borders plans to continue with community interventions designed to address the specific needs of her clientele.

Ms. Borders is currently serving as the regional education chair on the Health Advisory Board for East Tennessee, serves on the Loudon County Health Improvement Council and is president of the local Habitat for Humanity board of directors. Her involvement in the community and with her family motivate and inspire her daily.