

EFFECTS OF A RANDOMIZED TRIAL AFTER-SCHOOL PHYSICAL ACTIVITY CLUB  
ON THE MATH ACHIEVEMENT AND EXECUTIVE FUNCTIONING OF GIRLS

By

Katelin da Cruz

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

School Psychology - Doctor of Philosophy

2017

## ABSTRACT

### EFFECTS OF A RANDOMIZED TRIAL AFTER-SCHOOL PHYSICAL ACTIVITY CLUB ON THE MATH ACHIEVEMENT AND EXECUTIVE FUNCTIONING OF GIRLS

By

Katelin da Cruz

This study tested the hypothesis that an after school physical activity intervention would increase academic achievement in math and executive functioning in adolescent girls living in low-income communities. The study also examined how the change in executive functions relates to change in math achievement. This research was conducted using participants from the Girls on the Move (GOTM) Intervention project (Robbins et al, 2013) where participants were randomized to control or intervention at the school level. Girls in 5<sup>th</sup> to 8<sup>th</sup> grade participated in an after-school physical activity club three days per week for seventeen weeks. Control groups received their usual school offerings. All measures were administered at pre-test and post-test. Math achievement was measured by Woodcock-Johnson Test of Achievement-III Applied Problems and Math Fluency subtests. Executive function was measured by the National Institutes of Health Toolbox List Sorting Working Memory, Flanker, and Dimensional Card Sort tasks. MVPA was assessed with accelerometers. Cardiovascular fitness was assessed by estimating VO<sub>2</sub> max using PACER scores. Height and weight was assessed to calculate body mass index. Percent body fat was estimated using foot-to-foot bioelectric impedance models. Results from repeated measures ANOVA indicated that the intervention group showed more improvement in inhibition and math fluency performance than the control group. No mediation effects of change in executive function on the relation between intervention status and change in math achievements were found. Overall, these results indicate that physical activity intervention

can improve inhibition and math fluency but that change in executive function may not be the causal mechanism by which physical activity intervention affects math achievement.

## ACKNOWLEDGEMENTS

I want to thank my advisor, Dr. Jodene G. Fine, for her support and guidance throughout the graduate school and dissertation process. For accepting me as a graduate student advisee, expressing enthusiasm about my research interest, and working quickly to help me secure funding for this dissertation I will always be grateful. I also want to thank my other dissertation committee members, Drs. Karin Pfeifer, Matt Pontifex, and John Carlson, who provided advice and enthusiasm, even when they were out of the country. This project would have been only a dream without the support and collaboration of many individuals in the College of Nursing, particularly Dr. Lorraine B. Robbins, Stacey Wesolek, and Kelly Bourne, and my research assistants, Anthony Delli Paoli, Alexis Downes, Sarah Hackert, Nicole Mathes, Rachel Ogle, Olufemi Oluyedun, and Justina Yohannan. Thanks are also due to the data collection schools and their staff for their cooperation, flexibility, and frequent small kindnesses. Finally, I want to express my extreme gratitude to my family, friends, and cohort for their support and encouragement. This research was supported by the American Psychological Foundation's David H. and Beverly A. Barlow Grant and funds from Michigan State University.

## TABLE OF CONTENTS

LIST OF TABLES.....	vii
LIST OF FIGURES.....	ix
LIST OF ABBREVIATIONS.....	x
CHAPTER 1: INTRODUCTION.....	1
Purpose.....	1
Background.....	1
Importance.....	4
CHAPTER 2: LITERATURE REVIEW.....	7
Physical Activity Trends in School-Aged Youth.....	7
Health Indicators and Academic Achievement.....	11
Physical Activity and Academic Achievement.....	11
Obesity and Academic Achievement.....	18
Cardiovascular Fitness and Academic Achievement.....	21
Executive Functioning.....	22
Executive Functioning and Mathematics.....	25
Health Indicators and Executive Functioning.....	31
Physical Activity and Executive Functioning.....	31
Obesity and Executive Functioning.....	36
Cardiovascular Fitness and Executive Functioning.....	38
Proposed Biological Mechanisms.....	40
Purpose of the Present Study.....	46
Research Questions and Hypotheses.....	49
CHAPTER 3: METHODS.....	53
Study Approval and Funding.....	53
Trial Design.....	53
Study Procedures.....	54
Recruitment Procedures.....	54
Intervention Procedures.....	55
Treatment Fidelity Procedures.....	56
Data Collection Procedures.....	57
Participants.....	58
Measures.....	59
Math Achievement.....	59
School Attendance.....	60
Executive Functioning.....	60
Intervention Dosage.....	66

Physical Activity.....	66
Body Composition.....	69
Cardiovascular Fitness.....	70
Demographic Information.....	71
CHAPTER 4: RESULTS.....	73
Sample Characteristics.....	73
Missing Data.....	78
Intervention Fidelity.....	79
Changes in Cardiovascular Fitness, Body Composition, and Physical Activity.....	83
Research Question 1.....	86
Research Question 2.....	91
Research Question 3.....	96
CHAPTER 5: DISCUSSION.....	101
Results of Intervention on Physical Activity.....	102
Results of Intervention on Executive Functioning.....	103
Results of Intervention on Math Achievement.....	108
Limitations.....	113
Future Research.....	116
Implications.....	120
Concluding Remarks.....	122
APPENDICES.....	123
Appendix A. Parent and Student Consent Forms.....	124
Appendix B. Timeline for completion of study.....	128
Appendix C. Variable Correlations.....	129
REFERENCES.....	130

## LIST OF TABLES

Table 1 List of Constructs and Measures to be Used in Analysis of Research Questions.....	71
Table 2 Research Questions, Statistical Questions, Analysis, and Variables.....	72
Table 3 Descriptive Statistics for Categorical Demographic Variables.....	74
Table 4 Descriptive Statistics for Measures Prior to Maximum Likelihood Estimation of Missing Data.....	75
Table 5 District Demographic information from the Performance and Information Single Record Student Database (SRSD).....	76
Table 6 Crosstabulation of Racial Category and Intervention Group.....	77
Table 9 Crosstabulation of Free and Reduced Lunch Status and Intervention Group .....	77
Table 8 Descriptive Statistics of Intervention Fidelity Checklist Measures.....	80
Table 9 Results of Chi-square Test and Descriptive Statistics for Club Attendance Level by School.....	81
Table 10 Univariate Effects for School on Club Fidelity Measures.....	82
Table 11 Means and 95% Confidence Interval for groups at Pre and Post-test and Change Values.....	83
Table 12 Repeated measures ANOVAs for VO2 Max, MVPA, BMI, and Body Fat Percentage.....	84
Table 13 Means and 95% Confidence Interval of Executive Function Variables for Groups at Pre and Post-test and Change Values .....	87
Table 14 Repeated Measures ANOVAs for Working Memory, Inhibition, and Shifting.....	89
Table 15 ANCOVA table for Post-Test Math Fluency (WJMF) and Problem Solving (WJAP) by Intervention and Controlling for Pre-test WJMF and WJAP.....	92
Table 16 Repeated Measures ANOVAs for Math Fluency and Math Problem Solving.....	93
Table 17 Means and 95% Confidence Interval of Math Achievement Measures for Groups at Pre and Post-test and Change Values.....	95

Table 18 Pearson Correlations Between Executive Function and Math Achievement Change Scores.....	97
Table 19 Model Summary, Regression Coefficients, Standard Errors, and Significance Values for the Simultaneous Multiple Mediator Model of Math Problem Solving.....	98
Table 20 Model Summary, Regression Coefficients, Standard Errors, and Significance Values for the Simultaneous Multiple Mediator Model of Math Fluency.....	99
Table 21 Variable Correlations.....	129



## LIST OF FIGURES

Figure 1 Conceptual Model of Link Between Physical Activity, Executive Function, and Academic Achievement for Current Study.....	3
Figure 2 CONSORT Diagram of Loss and Exclusions.....	73
Figure 3 Profile Plot of Intervention and Control Group Marginal Means on NIH FIT at Pre-test and Post-test.....	90
Figure 4 Profile Plot of Intervention and Control Group Marginal Means on Math Fluency (WJMF) at Pre-test and Post-test.....	94
Figure 5 Regression Coefficients for Simultaneous Multiple Mediation Model of Problem Solving.....	98
Figure 6 Regression Coefficients for Simultaneous Multiple Mediation Model of Math Fluency.....	100

## LIST OF ABBREVIATIONS

CVF	Cardiovascular fitness
EF	Executive Functioning
MVPA	Moderate to Vigorous Physical Activity
PA	Physical Activity
WJ-III	Woodcock-Johnson Test of Achievement Third Edition
WJMF	Woodcock-Johnson Test of Achievement- Math Fluency
WJAP	Woodcock-Johnson Test of Achievement- Applied Problems
NIH DCCS	NIH Toolbox- Dimensional Change Card Sort
NIH FIT	NIH Toolbox- Flanker Inhibitory Test
NIH LS	NIH Toolbox- List Sorting

## **CHAPTER 1: INTRODUCTION**

### **Purpose**

The main objective of this study was to evaluate if an after-school physical activity club improves academic achievement in math and executive functioning in adolescent girls who live in low-income communities. In addition, relations between math achievement, executive function, and physical activity were explored to help understand potential causal mechanism. Physical activity can affect the brain's physiology in many ways. These physiological changes may cause improved cognitive functioning. In particular, cognitive processes termed executive functioning may be positively affected. Executive functioning are cognitive processes required to successfully complete actions that lead to a specific goal (Diamond, 2013). It has been hypothesized that this improved executive functioning may also influence academic performance (Tompsonowski, Davis, Miller, & Naglieri, 2008). This research study used a randomized control trial design to evaluate if participants in the after-school club designed to promote physical activity showed stronger executive functioning and improved math achievement compared to participants in the control group.

### **Background**

Regular participation in physical activity is an acknowledged part of a healthy lifestyle (Department of Health and Human Services and Department of Education, 2000; U.S. Department of Health and Human Services, 2008). However, research has indicated that current generations of children are less active than previous generations (Boreham & Riddoch, 2001; Burkhalter & Hillman, 2011). As students have become less active outside of school there has

also been a decrease in the opportunities for students to be physically active while at school. The number of students attending physical education (PE) classes daily decreased from 41.6% of students in 1991 to 25% of students in 1995 and has stayed at these lower levels (29% in 2013; (Frieden, Jaffe, Cono, Richards, & Iademarco, 2014).

These trends continue despite research showing that educational policies requiring increased opportunities for physical activity during the school day can be effective in promoting regular physical activity (Institute of Medicine, 2013). Although schools are a convenient venue for helping youth to meet physical activity recommendations, administrators face barriers to devoting time in the school day to physical activity. Researchers have found that a combination of priority for standardized testing and insufficient resources are cited by educators and policy makers as barriers to successful implementation of strong policies on physical activity in schools (Amis, Wright, Dyson, Vardaman, & Ferry, 2012). There is increased pressure on schools to improve academic achievement due to the No Child Left Behind Act of 2001 (Center for Educational Policy, 2007). Those who are under pressure to improve test scores may believe that time devoted to non-academic activities will negatively influence children's academic skill development (Lindner, 2002; Morales et al., 2011) and therefore result in lower performance on standardized achievement tests that are reported in line with the regulations of the No Child Left Behind Act (NCLB, 2002).

However, this concern regarding the effects of instructional time lost to physical activity may not properly take into account benefits of physical activity for academic achievement. Beyond the known positive health effects, researchers have begun to suggest that physical activity may have beneficial influences on cognitive and academic functioning (Tomporowski et al., 2008; U.S. Department of Health and Human Services, 2010). This research draws on the executive

function hypothesis, which predicts that the largest improvements in cognition due to aerobic exercise will be in executive functioning (Hall, Smith, & Keele, 2001). Strong support for this hypothesis has been found in work on preventing cognitive decline in older adults (Colcombe & Kramer, 2003). More recently, the possibility that this may be true for the developing cognitive function of youth has been raised by promising research (Tomporowski, Lambourne, & Okumura, 2011). Children's executive functioning may be sensitive to the influence of physical activity because over the course of childhood and adolescence, executive functioning improves and these improvements parallel structural and functional development of the brain (Bunge & Crone, 2009). Correlational research showing positive associations between physical activity or participation in physical education and cognitive functioning provides preliminary support that this executive function hypothesis may also be true for children (Fedewa & Soyeon, 2011). Executive functioning includes higher level cognitive abilities such as holding knowledge in working memory, shifting between strategies to effectively plan, and to inhibit impulsive behaviors. As will be discussed further, theory and some preliminary research indicate these skills support academic achievement. Therefore, exercise induced improvements in executive functioning may also lead to improvements in academic performance and academic behavior, as outlined in Figure 2 below.

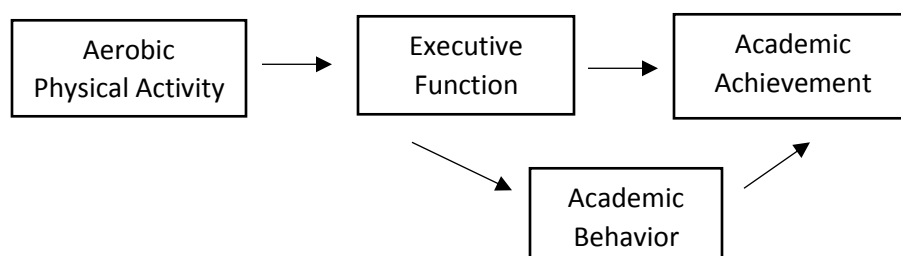


Figure 1 Conceptual Model of Link Between Physical Activity, Executive Function, and Academic Achievement for Current Study

The combination of these trends has prompted renewed interest in examining the links between physical activity and cognition and academic achievement in children. A historical review shows that there were a large number of studies on the psychological benefits of physical activity in the 1950s and 1960s but that interest in this line of research declined in the 1970s and 1980s (Kirkendall, 1986). When interest in the area was renewed in the 1990s, it was focused on the effect of activity on aging with few studies on how physical activity influences children's development (Tompsonowski et al., 2008). Therefore, currently the research on the effects of activity on children's cognitive development and academic achievement is primarily associational, with only a few randomized control trials.

### **Importance**

While schools are a perfect venue for helping all youth to be physically active, school districts do not appear currently to be prioritizing physical activity. In the United States, elementary and secondary school students spend about half of their day at school or going to and from school. It is unsurprising therefore that recent analysis of school policy changes suggests that policies supporting physical activity could significantly help students meet government guidelines. For example, policies adding mandatory daily physical education class have been found to add about 23 minutes of moderate to vigorous physical activity to a student's day (Bassett et al., 2013). This means there is a potential for daily physical education to help students achieve one-third of the 60 minutes of moderate to vigorous physical activity recommended (U.S. Department of Health and Human Services, 2008b). After-school activity programs have the potential to add another ten minutes (Bassett et al., 2013). However, levels of physical education in schools remains low and close to half of youth in grades 9-12 do not meet government physical activity guidelines (Centers for Disease Control, 2014). A lack of clear

evidence about the benefits of physical activity for school-aged children's academic achievement may be causing many stakeholders to miss an opportunity to improve cognitive functioning and academic performance, while also supporting physical health, by providing youth with appropriate amounts of physical activity.

Physical activity could provide an additional venue for promoting academic success that would also provide physical health benefits. This study will help clarify the existing body of literature regarding these possible benefits of physical activity. If successful, this study will provide evidence, using a strong design, regarding the efficacy of physical activity to promote cognitive health and academic growth of students. In addition, the exploration of the relations between physical activity and fitness, executive function, and math achievement will improve the understanding of these connections. Together, this may inform the strategies and tools educators and public health officials utilize to support executive functioning and academic achievement.

In addition, this study has particular importance because of the population involved. Adolescent girls in urban, low-income schools (the population for this study) have an increased likelihood for poor physical health and academic outcomes. First, in the area of physical health, a much higher percentage of adolescent males participate in vigorous physical activity than do their female peers. This is true within all racial and ethnic groups and for all high school grade levels. Also, females in families of low socioeconomic status (SES) tend to have higher rates of obesity (Ogden, Flegal, Carroll, & Johnson, 2002). In addition, the decline in physical activity seen overall in adolescence is greater among urban, minority girls of low SES. Therefore, the population in this study is at increased risk for poor physical activity and obesity outcomes. Minority status, SES, and school location all continue to be strongly associated with academic achievement in the United States (Sirin, 2005). Many studies indicate that SES is an important

predictor of neurocognitive development, particularly executive function (Hackman & Farah, 2009). Therefore, this population of female, minority, and low SES students would particularly benefit from interventions aimed at physical health, executive functioning, and academic performance.



## **CHAPTER 2: LITERATURE REVIEW**

The following chapter reviews research literature on the connections between health outcomes, executive functions, and math achievement in order to understand why a physical activity intervention may be hypothesized to influence executive functions and math achievement. An overview of current physical activity levels in youth, with an analysis of trends in adolescent minority females is included first. Next, the existing literature on the relations of physical activity, obesity, and cardiovascular fitness with math achievement and executive functions is explored to provide an understanding of how the current study adds to the existing knowledge in this area. A brief discussion of proposed biological mechanisms for these connections is provided to add to an understanding of why these relations may exist. The chapter concludes with the purpose of the present study, pulling together the literature to inform the study's research questions and hypotheses.

### **Physical Activity Trends in School-Aged Youth**

Researchers suggest that physical activity (PA) levels among children and adolescents have shown significant declines over the last several decades while engagement in sedentary behaviors has escalated (Brownson, Boehmer, & Luke, 2005; Pate, Mitchell, Byun, & Dowda, 2011). Brownson and colleagues (2005) found relatively stable levels of leisure-time PA, declining use of walking and biking for transportation and activity in the home, and increasing engagement in sedentary activities combine to result in an overall trend of declining total PA. The *Physical Activity Guidelines for Americans* were published in 2008 amidst these concerns that children and adolescents in the United States are not engaging in enough PA (U.S. Department of Health and Human Services, 2008). They provide PA guidelines specifically for children and adolescents based on a comprehensive review of PA and health research regarding

the level of PA needed to achieve the desired health benefits. It is recommended that children and adolescents engage in 60 minutes or more of PA daily and that most of the 60 minutes a day should be either moderate- or vigorous-intensity aerobic physical activity. However, according to data from the 2013 Youth Risk Behavior Surveillance System (YBRSS), only 27% of American youth in grades 9-12 met the *Physical Activity Guidelines for Americans* described above (Centers for Disease Control, 2012). The data from Michigan youth (26.7%) is similar to the rates reported for the nation as a whole.

This trend in youth is likely due to social and environmental factors, one of which may be decreased time for PA in schools and during afterschool programs. The environment that schools provide plays a critical role in the physical activity behaviors of children and adolescents. In particular, physical education (PE) classes have the potential to provide educational experiences that assist youth in leading a healthy, active lifestyle. However, opportunities for youth to regularly engage in PA throughout the school day appear to be increasingly limited. The Centers for Disease Control (2014) recommends that school health programs provide students with a substantial portion of the recommended daily amount of PA through physical education class. However, according to the YBRSS, only 29.5% of American students in grades 9-12 in 2013 reported having daily physical education (Centers for Disease Control, 2014). This is unsurprising, given that according to the national School Health Policies and Practices study only 2% of middle schools in the United States required daily physical education during the entire school year (Lee, Burgeson, Fulton, & Spain, 2007) In addition, research into the amount of time spent in moderate-to-vigorous physical activity (MVPA) shows that many physical education classes may be falling short of providing students with a good percentage of the recommend PA. Meta-analyses of studies conducted in both elementary (Fairclough & Stratton, 2006) and middle

and high schools (Fairclough & Stratton, 2005) have found that students were engaged in MVPA less than half their PE class time. Therefore, current physical education policies and practices may be acting as barriers to fulfillment of the governmental PA guidelines.

Recess and after-school programs are also opportunities to offer PA in school settings to our school aged youth. As of the 2011-2012 school year, only 10% of states in the U.S. required daily recess at the elementary school level (Institute of Medicine, 2013). Analysis of policy changes found that enhanced-quality recess (with age-appropriate fitness equipment, blacktop games, and adult encouragement) can increase the time children spend being physically active during recess from an average of nine minutes to fourteen minutes (Bassett et al., 2013). While not universal in potential reach like recess, afterschool programs do reach many school-aged youth in the United States. Reviews of the effects of afterschool programs designed to promote physical activity and/or fitness have determined that these types of afterschool programs result in numerous health benefits (Beets, Beighle, Erwin, & Huberty, 2009; Pate & O'Neill, 2009). However, although these after-school programs provide additional venues for supporting PA, researchers have found that there is limited focus on providing physical activity through these venues. For example, only 30% of states organizations related to after-school program quality have policy documentation addressing the provision of opportunities for PA during these programs (Beets, Wallner, & Beighle, 2010). If, as this information suggests, current school environments and policies are not set up to support the levels of student PA needed for the known health benefits of exercise they are likely not set up to support the potential benefits for academic performance.

Researchers have identified many correlates of PA. Of particular importance for this study are age, sex, socioeconomic status, and ethnicity (Institute of Medicine, 2013). Overall, PA

levels tend to decline from infancy through adolescence and adulthood. For example, a study in Massachusetts found that while almost all of the students in grades 1-6 performed at least 60 minutes of MVPA at least five days a week, fewer than one-third of high school students did (Pate et al., 2002). Physical education requirements, and actual student participation rates, decline substantially between 8th and 12th grades (Johnston, Delva, & O'Malley, 2007). Johnston and colleagues found that in a nationally representative sample, about 87% of 8th graders were in schools that had a physical education requirement, compared to only 20% of 12th graders.

This decline in physical activity seen during adolescence is evident among adolescent girls and particularly adolescent girls of minority and low SES families. In correlational research PA level has been found to be associated with both SES and gender (Molnar, Gortmaker, Bull, & Buka, 2004). First, a decline in PA levels are seen in girls overall as they move towards adolescence (Centers for Disease Control, 2012). According to the 2003-2004 National Health and Nutritional Examination Survey, by 9<sup>th</sup> grade only around 22% of girls meet the national recommendations for daily PA (Troiano et al., 2008). For adolescent girls educated in low SES schools, the school day is unlikely to provide opportunities for PA. High SES schools have been found to be more likely to require physical education, to have more frequent physical education classes, to have students spend more time in vigorous exercise in those classes, and to have participation in more PA outside of school (Johnston et al., 2007; Sallis, Zakarian, Hovell, & Hofstetter, 1996). Because of this increased risk, it is important to identify physical activity interventions that can create long-lasting changes in the activity level of this population.

Adolescent girls in low-income and high-minority population schools would benefit from support in increasing physical activity and from the potential improvements in executive

functioning and academic performance. This population is at increased likelihood for poor academic outcomes. Socioeconomic status, minority status, and school location in a low-income community all continue to be strongly associated with academic achievement in the United States (Sirin, 2005). Many studies indicate that socioeconomic status is an important predictor of neurocognitive development, particularly executive functioning (Hackman & Farah, 2009). Taken with the increased likelihood of decreased physical activity in this population, female youth living in high minority, low-income communities are an appropriate population for the current study.

## **Health Indicators and Academic Achievement**

### **Physical Activity and Academic Achievement**

It is well accepted that a moderate amount of physical activity can have significant health benefits, such as reduced risk for cardiovascular disease and diabetes and overall improvement health of muscles, bones, and joints (U.S. Department of Health and Human Services et al., 1996). However, researchers have begun to also examine the influence of differing levels of physical activity on academic achievement. A wide variety of studies examining physical activity and academic achievement have been conducted, using various measures of academic performance. The effects of regular physical activity on academic performance have primarily been shown to be neutral or positive. However, the extant studies vary significantly in intensity and duration of physical activity and the area of research overall has suffered from the level of methodological quality. Because of these limitations, the proposed link between physical activity and academic success may not have been conclusively assessed.

Although multiple correlational studies have been conducted to examine the relation between physical activity and academic achievement, the findings from these studies are not conclusive. Five out of the seven large scale correlational studies identified found at least modest support for a positive relation (Dwyer, Sallis, Blizzard, Lazarus, & Dean, 2001; Fox, Barr-Anderson, Neumark-Sztainer, & Wall, 2010; Kantomaa, Tammelin, Demakakos, Ebeling, & Taanila, 2010; Nelson & Gordon-Larsen, 2006; Sigfusdottir, Kristjansson, & Allegrante, 2007). Research involving almost 8,000 children between the ages of seven and nine in Australia found modest though significant correlations between ratings of scholastic ability by teachers and self-reported physical activity in the previous week for both genders and all age groups. (Dwyer et al., 2001) Nelson and Gordon-Larsen (2006) found that more active adolescents were more likely to have higher grades in math and science. Fox and colleagues (2010) found a significant linear association between performing more hours of MVPA and a higher mean grade point average (GPA) for both gender middle school students. In a large correlational study from Finland, adolescents self-report of physical activity level was related to self-reported academic performance for both genders, with the group of highly physically active adolescents twice as likely to report high overall academic performance (Kantomaa et al., 2010). In a sample of 14 and 15 year-old Icelandic students, there was a significant positive relationship between self-reported physical activity and core subject grades (Sigfusdottir et al., 2007). However, it was of very modest strength and was a weak predictor when models controlled for other variables. Overall, the majority of studies have found a relation and those correlational studies have shown that relation to be modest.

In addition to the modest strength of the relation found in the studies above, two correlational studies have found no significant relation between measures of physical activity and

academic performance. It is possible that the variability in the findings in this area is due to the use of children's self-reported physical activity, which is known to lack reliability (Sirard & Pate, 2001). All the correlational studies of physical activity identified relied on self-report of physical activity. In one of the two studies finding no significant relations, researchers examined correlations between end of the year examination results in English, math, and science and a self-report of minutes and frequency of physical activity (Daley & Ryan, 2000). In the other study, Yu and colleagues also found no significant relation between self-report physical activity and test scores in a sample of Chinese school children between the ages of eight and twelve (Yu, Chan, Cheng, Sung, & Hau, 2006). These studies, in combination with the modest relationships found in the studies with positive results suggest that further research utilizing different methods examining this area is necessary.

An additional research method that has been used to examine this question is the use of longitudinal information. Using a longitudinal dataset, Stevens and colleagues (2008) examined the importance of physical activity and physical education in the prediction of academic achievement. Structural equation modeling techniques indicated that for boys and girls in fifth grade, parent-reported physical activity was significantly related to both math and reading achievement. Physical education participation was not related. However, as was seen with the correlational research previously noted, the measurement of physical activity was limited. In this dataset, the questions on physical education were ordinal (e.g. how many days do your students receive physical education class) and didn't provide any information about how much moderate to vigorous physical activity students received while in physical education class. More detailed and objective information about children's physical activity, such as minutes spent physically engaged and intensity of exercise, would have provided higher quality data. The restriction in

variability and lack of reliability in reporting could have influenced the size of the association with academic achievement.

Many reviewers have commented on the need for research in authentic settings in order to understand the influence on academic achievement when students are given increased physical activity opportunities in schools (Lees & Hopkins, 2013; Singh, Uijtdewilligen, Twisk, van Mechelen, & Chinapaw, 2012; Tomporowski et al., 2008). The well-known *Trois Rivieres* study is an early randomized study where researchers conducted an experiment in an authentic school setting (Shephard, 1996). They provided an intervention group with an additional twenty minutes of physical education a week as students moved through grade 1 to grade 6 in Quebec. While students in the control group initially had higher grades, after the intervention began the experimental group began to outperform the control group overall in academic performance. However, it should be noted that despite a strong experimental design, the study suffers from the limitation of using teacher-assigned academic grades as the primary outcome measure despite the possibility that changes might be due to teacher expectancies or response bias. Thus, although the study produces a strong result from a well-accepted experimental design, the outcome measures used may not be reliable. In an early experimental study in this area, children in intervention schools were provided a physical activity program titled Sports, Play, and Active Recreation of Kids (SPARK) for three days a week the entire school year while children in the control continued with their typical physical education curriculum (Sallis et al., 1999). Although all of the students in the study experienced declines in their performance on the Metropolitan Achievement Tests (Psychological Testing Corporation, 1990), the intervention group showed significantly less of a decline than the control group. Dwyer and colleagues (2001) created a control group, a skills group that had higher duration and frequency of PE, and a fitness group



who had the same duration and frequency of PE as the skills group but with a focus on raising the heart rate. There were no significant group differences in mathematics or reading performance, as measured by national standardized tests. This evidence does not support a positive effect of enhancing existing physical education time. However, interpretation is challenging because non-scientifically validated standardized tests were used as outcome measures in the study that did not find group differences (Dwyer et al., 2001). Overall, the majority demonstrate only that time in physical activity is not detrimental to student academic performance.

Researchers who have recently examined models other than modifying physical education have found evidence for specific academic area improvements. One study examined a program combining 5-20-minute classroom physical activities modified from the Take 10! Program with structured recess aimed at promoting physical activity. Math problem-solving scores but not reading comprehension scores of third and fourth grade students in the intervention schools improved significantly more than students in the control condition. Students in the intervention group that were rated by teachers as having poor adaptive skills improved both their reading and math scores at a significantly greater rate compared to the students with similarly low ratings in the control group (Murray et al., 2008). In a second study, 90 minutes per week of physically active academic lessons were promoted in second and third grade classrooms (Donnelly et al., 2009). The changes seen in the intervention group on a norm-referenced academic measure of math, reading, spelling, and the achievement composite (Wechsler Individual Achievement Test-2nd Edition; The Psychological Corporation, 2001) were significantly greater than the control group. In a similar physical activity program to that completed by Donnelly and colleagues (2009), researchers completed a comprehensive

promotion of physical activity (in classrooms, during recess, in the hallways, etc.) called the Action Schools! BC model in cluster randomized intervention schools. This program created an average of 47 additional school day minutes of PA per week (Ahamed et al., 2007). There was no significant difference in total score on the Canadian Achievement Test 3 (CAT 3) at follow-up. However, it should be noted that this research only analyzed overall academic performance instead of looking at the influence on specific subject areas or skills (e.g. reading comprehension or math problem solving) and again did not utilize a scientifically validated and norm-referenced measure of academic achievement. Therefore, there is a difference in the outcomes seen by the academic achievement measures used and if they were standardized and or assessed separate skills. Additional studies that also allow for more differentiated examination of academic outcomes and utilize rigorous academic achievement measures are needed to improve understanding of this relation.

Another option for increasing PA opportunities, which is utilized in the current study, is the use of an after-school program. In a study using this model, Davis and colleagues (2007) used an individual level randomized control design to examine academic performance following an afterschool physical activity intervention. This study randomly assigned children between the ages of seven and eleven to aerobic exercise of a low dose (20 minutes per day) or high dose (40 minutes per day) or to a no exercise control for thirteen weeks. There was a difference in performance between both exercise groups and the control group on math performance, but no difference between the high and low dose groups and no differences between any groups on reading performance. However, this study was completed using a sample that was overweight (>85<sup>th</sup> percentile Body Mass Index) in addition to inactive. Therefore, while this research

addresses the need for more rigorous research in this area, it is not generalizable to non-active but non-overweight populations.

In summary, although there is some evidence for the relation between academic achievement and physical activity, overall the correlational and experimental research in physical activity and academic performance is mixed. Correlational research continues to be the norm in this area of work and the limited experimental work currently available either suffers from limitations of the outcomes measures used or generalizability to a non-overweight population. However, no cross-sectional or experimental work has determined that there are negative influences of improved physical education or increased PA opportunities on academic performance. Overall, the results of existing studies vary between suggesting no effect to suggesting a moderate effect. Possible explanations for this variation include the type and specificity of academic outcome measures used, the nature of the physical activity in which students in the program engaged, and population characteristics.

These physical activity interventions may result in students engaging in more physical activity, which could have an effect on their cardiovascular fitness (CVF) and/or body composition. However, the studies presented above do not explicitly explore the relation of any changes in cardiovascular fitness and body composition as they relate to the changes in academic performance measured as a result of the intervention. A recent review argues that researchers have demonstrated that overweight children perform less well on math and reading assessments, even after adjusted for SES and physical activity (Burkhalter & Hillman, 2011). Therefore, these related health measures should be considered.

## **Obesity and Academic Achievement**

It is possible that the effects of being overweight are felt in various aspects of school-life, including academic achievement. There are multiple pathways through which this effect might transfer. It has been established that overweight and obese children are at higher risk for low self-esteem and that they have higher rates of anxiety disorders, depression, and other psychopathology (Taras & Potts-Datema, 2005) which may make it harder for children to be attentive in class or increase school refusal, thereby preventing them from learning. This psychosocial aspect has been emphasized in the literature. In addition, health problems associated with being overweight may increase sick days, leading to missed academic time that subsequently affect school performance. One study by researchers at the National Center for Chronic Disease Prevention and Health Promotion of the CDC found that overweight and obese adolescents had 36% and 37% more missed school days, respectively, than adolescents of normal weight. In addition, obese adolescents had a significantly higher likelihood of missing greater than 4 days of school per year than adolescents of normal weight (Pan, Sherry, Park, & Blanck, 2013). The combination of higher rates of anxiety and depression and increased absences from school in this population provides a theoretical reason that obesity may be linked to academic achievement.

There is a small body of research on the association of measures of body composition (Body Mass Index and body fat percentage) and academic performance. In one study using a large data set of self-report data of Icelandic adolescents, the correlation between Body Mass Index (BMI) and grades was significant in a negative direction but of modest strength ( $r = -0.12$ ). It was consistent with the hypothesis that students with higher BMI would show lower academic achievement. (Sigfúsdóttir, Kristjánsson, & Allegrante, 2007) However, while BMI

was strongly associated with academic achievement it was overshadowed by the relation of parental education, school absenteeism, and self-esteem when included together in a model. Therefore, the effect of overweight on academic achievement may indeed be through school absenteeism and self-esteem. While the study sample was large, the data collection measures relied on self-reports from adolescents, including a composite measure of academic achievement based on self-reported grades. Several studies of adolescent populations have examined the validity of self-reports of height and weight and have found significant under-reporting of weight and significant over-reporting of height (Strauss, 1999; Wang, Patterson, & Hills, 2002). As reviewed earlier, self-report of physical activity has significant limitations of bias and variability (Sirad & Pate, 2001). Authors of a meta-analysis of self-reported grades, class rank, and GPA concluded that self-reported grades differ in their validity by participant school performance and cognitive ability and therefore school be used with caution (Kuncel, Crede, & Thomas, 2005). Research by the ACT shows that the median percentage of high school students who self-reported the same grades as shown on their transcript was 68% (Sanchez & Buddin, 2016). Therefore, the findings should be interpreted with caution due to the possibility that students did not accurately report their grades, weight, or level of physical activity.

However, more current research using more rigorous methods has found similar relations. Research completed in the Texas schools using direct data found similar results as this self-report analysis. The prevalence of students meeting standards on the state achievement test (TAKS) was significantly higher in the lowest BMI group compared to all other categories, regardless of sex or grade (Janak et al., 2014). In addition, Janak and colleagues found a significant association of BMI with academic achievement, although the relation was weaker after adjusting for socioeconomic status. The linear modeling suggested a 5% increase in the prevalence of

students meeting healthy BMI standards would result in a 2.25% increase in the prevalence of students meeting the TAKS standard. In the study by Kamijo and colleagues (2012), preadolescent children between 7 and 9 years completed the Wide Range Achievement Test (WRAT3; Wilkinson, 1993). Regression analyses indicated that higher BMI was associated with lower spelling and math achievement but not with reading achievement. The regression analyses for body fat indicated that a higher percentage of whole body fat was associated with lower reading and spelling achievement. Therefore, both measures of body composition and obesity indicated relationships with spelling while the different measures were related either with reading or math achievement. Together this data suggests that measures of obesity such as BMI are related to academic performance when using standardized achievement measures as well as more rigorous body composition measures rather than self-report. However, the differences in academic area indicated between the body fat measure and the BMI measure require further exploration.

Researchers utilizing the ECLS-K longitudinal data set have been able to examine this question using cross-sectional methods as done in the studies above as well as longitudinal methods. Reading and math test scores at baseline are significantly higher among never overweight children compared to the other two groups (Datar & Sturm, 2006). Children who were overweight currently or in the past had significantly more school absences in kindergarten and third grade and were also more likely to repeat grades. Of particular relevance to this study which is completed using a female only sample, these researchers also found that moving from not-overweight to overweight status between kindergarten entry and end of third grade was significantly associated with reductions in test scores and teacher ratings of approaches to learning among girls. However, this link was not seen in the boys in the sample. Therefore,

change in overweight status during the first four years in school does appear to be a significant risk factor for poor school outcomes, particularly among girls.

The information currently available does suggest a relation between body composition or obesity and overall academic performance as well as other academic outcomes like attendance. However, little information is available about how this relation may differ by academic skill. The one study that examined the relation to different academic areas found different academic and obesity associations based on the body composition measure (BMI or body fat percentage with DXA scan). Body composition has also frequently been shown to have a significant negative correlation with cardiovascular fitness in youth, another indicator of overall health and fitness. However, as cardiovascular fitness is another significant aspect of health and has been examined in relation to academic achievement and executive function, its relations to these areas will be discussed separately below.

### **Cardiovascular Fitness and Academic Achievement**

There is less information is available about how cardiovascular fitness (CVF) is or is not related to academic achievement and differentially related to different academic skills. Castelli and colleagues (2007) assessed the cross-sectional association between CVF (using the PACER measure used in the current study) and state-wide achievement tests (Illinois State Achievement Test) in 3rd through 5th grade students. There was a positive and similarly strong association between CVF and total achievement ( $r = 0.48$ ), math ( $r = 0.49$ ) and reading ( $r = 0.45$ ). Similarly, significant associations with Texas state-wide academic achievement tests (TAKS) in elementary through high schools have been reported, also using the PACER measure (Janak et al., 2014; Welk et al., 2010). Janak and colleagues found a significant association of CVF and academic

achievement after adjusting for BMI and socioeconomic status. Linear modeling suggested a 5% increase in the prevalence of students meeting healthy CVF standards would result in a 0.65% increase in the prevalence of students meeting the TAKS standard. However, CVF was a poorer predictor of academic achievement and had a smaller effect size than BMI, which was found to predict an increase of 2.25% of student meeting the TAKS standard. In addition, the prevalence of students meeting standards on the state achievement test (TAKS) was significantly higher in group with the highest CVF scores compared to all other categories, regardless of sex or grade (Janak et al., 2014). Researchers in Japan found that time on a one-mile run/walk test added significantly to the prediction of GPA ( $P < 0.01$ ) in adolescent boys while measures of strength did not (Kalantari & Esmailzadeh, 2016).

Together, this suggests that there is an association between CVF and overall academic achievement. No information is available about how CVF is or is not differentially related to different academic skills. In addition, this relation may not be as strong as the relation between obesity and academic achievement. However, because research suggests that there are potential positive effects of improved cardiovascular fitness and body composition on academic skills, it is important to examine why these links exist. To explain this potential benefit of PA and improved body composition and CVF on academic performance, researchers have suggested improvements in executive functioning.

### **Executive Functioning**

To explain the potential benefit of PA and improved body composition and CVF on academic performance explored previously, researchers have suggested improvements in executive functioning. Theoretically, improvements in executive functioning as a result of increased physical activity may be part of the explanatory link between physical activity and



academic performance. In other words, a causal path may exist from physical activity to executive functioning to academic achievement. Therefore, in addition to examining academic performance as an outcome measure, researchers have recently begun to examine the potential of PA to support executive function development in youth. However, it is important first to determine if a relation between executive functioning and academic achievement has been determined.

Cognition is a general term while executive function (EF) is an umbrella term for specific cognitive processes involved in organizing and controlling goal-directed behavior (Banich, 2009). It is important to understand that executive functioning is multidimensional and complex. Because of this, there are a variety of models and viewpoints as to what components or subprocesses comprise executive functioning. Some of the components or processes hypothesized to fall under the category of executive functioning include: prioritizing behavior, inhibiting familiar behaviors, creating and maintaining an idea of what information is relevant for a current task, being resistant to information that is distracting or not relevant, switching between goals as things change, using relevant information in support of decision making, and handling novel information or new situations. As this suggests, the functions that fall under the category of executive function are indeed wide ranging.

In this analysis of the literature, executive functioning will be conceptualized as those components of cognition that involved in organizing goal-directed actions where those specific components can be somewhat differentiated by experimental tasks. There is a consensus that executive functioning is several underlying processes rather than a unitary construct. Although still debated in the field, a frequently accepted framework developed through latent analysis suggests that executive functioning consist of inhibition, updating or working memory, and

shifting (Miyake, 2000). Therefore, executive functioning in this review will look at these underlying processes but will not include research into general intelligence, simple information processing, or perceptual-motor tasks.

Inhibition is the control of attention, behaviors, and thoughts that are automatic but are not accurate or appropriate (Diamond, 2013). When a person exhibits inhibition, they are able to selectively focus on what they choose and to suppress attention to non-salient stimuli. Being able to stop impulsive actions or stay on task despite distractions (self-control or self-regulation) is strongly related to inhibition. Therefore, inhibition is thought to be important for behaviors that lead to success in the classroom. For example, inhibition may be important for being able to follow directions for assignments and controlling what one is paying attention to during assignments or lectures.

The second core executive function is working memory or updating. This executive function involves keeping information in memory and then doing some mental transformations on that information. Working memory is necessary for tasks that require someone to remember an event or information from earlier in time and relate it to information that comes later. For example, being able to remember the directions someone gave you to drive somewhere and then reverse those directions to drive home would require working memory. In relation to academic performance, this skill is thought to be necessary for making sense of written or spoken language, doing mental math, and incorporating new information when doing a math word problem (Diamond, 2013). It is important to remember that this is different from short term memory, which is holding information in memory without manipulating it. For example, repeating a phone number someone read aloud to you would require short term memory but not working memory, while repeating it backwards would involve working memory. In addition, the

construct of working memory is typically separated into two categories by the type of information being kept in memory (visual or verbal). Therefore, researchers must make decisions about which aspects of working memory will be examined.

The third executive function is alternatively called shifting or cognitive flexibility. This is the ability to adjust to changed demands, rules, or priorities in a situation. A frequently examined type of switching involves being able to change what dimension of a stimuli you are paying attention to, such as color or shape. Children generally are not able to succeed on this type of task until close to five years of age after which there is still variability in speed and accuracy. In relation to academic performance, the skill of shifting can be important for changing an approach to solving a math problem or changing your understanding of a sentence (Diamond, 2013). Research on the specifics of how these executive functions may influence math achievement is reviewed below.

### **Executive Functioning and Mathematics**

Many factors contribute to differences in mathematics achievement, such as domain-specific numerical skills and knowledge but also attitudes, motivation, language ability, intelligence, educational opportunity, etc. Other cognitive factors, such as executive functions, may also play an important role. Overall, the existing literature suggests that executive functioning influence academic performance (Mulder & Cragg, 2014). This may be an influence on a child's ability to apply appropriate knowledge or on his or her ability to display what they know in testing situations Being able to inhibit distractions or automatic responses (inhibition), being able to keep information in mind and manipulate it (working memory), and being able to switch strategies or rules (shifting) will all theoretically support a student's best performance (Diamond, 2013) Tomporowski et al., 2008). For example, in academic performance situations

(e.g. tests, being asked a question in front of the whole class) all of these three executive functioning theoretically would support a student in performing to his or her best ability. In addition, executive functioning is hypothesized to be required for many classroom behaviors that will allow students to benefit from instruction. Being able to inhibit impulsive behaviors is important for classroom success and as children move into higher grades, significant levels of inhibition are required during lectures. Theoretically, students need to utilize working memory resources efficiently to facilitate comprehension of complex material in the classroom or to solve problems using information they are given. Overall, strengths in executive functioning have the potential to improve academic performance while deficits may inhibit a student's ability to perform in the classroom and to show their knowledge.

In particular, research indicates a potential relation between executive functioning and mathematics achievement. Meta-analyses have suggested that children with mathematics disabilities have particular difficulty with the working memory, especially when numerical information is involved in the task (Andersson & Lyxell, 2007; David, 2012). In addition, correlational work using a range of measures has demonstrated that working memory uniquely predicts math achievement across a wide range of different ages (Cragg & Gilmore, 2014). Although less extensively studied and theorized about, most studies in the area suggest that inhibitory control abilities do related significantly to math achievement. Lastly, a meta-analysis of 18 studies on the relation of shifting ability and performance in math found an overall significant effect ( $r=.26$ ) (Yeniad, Malda, Mesman, van IJzendoorn, & Pieper, 2013)

However, mathematics is a broad subject area that includes diverse skills such as number sense, mathematical fact retrieval, math computation (addition, subtraction, multiplication, and division), geometry, and math problem solving or word problems. Thus, which component of EF

is involved may depend on the type of mathematical skill or ability that is considered, and on the strategies children use to solve the arithmetic problem. While the studies described above found relationships between executive function and math achievement they utilized single broad measure of math achievement that confound different math skills such as factual knowledge, math fluency, calculation or procedural knowledge, and math problem solving. Therefore, it is unclear from this information how these executive functioning may relate to specific areas of math skills (e.g. calculation or problem solving or math fluency). However, the fact that individuals can differ in their performance in these various aspects of math performance suggests that different components of math may rely with different weights on different sets of executive functioning skills. In support of this idea, there is emerging evidence that the contribution of executive skills may differ across these different mathematical skills.

**Executive functioning and math problem solving.** Researchers have suggested that executive function is more strongly related to measures of applied problem-solving (e.g. word problems) than the basic calculation skills emphasized in early elementary school (Best, Miller, & Naglieri, 2011). Solving mathematical word problems is a complex academic task involving multiple phases of problem solving, including establishing a problem representation, developing a solution plan, and executing each part of the solution plan while holding the plan in memory. Therefore, a general theory in this area is that working memory is important for handling each part of these different phases and for following through each phase. For example, the different information and relationships in the problem must be integrated into a problem representation, which requires holding each piece of information and relationship simultaneously in working memory (Swanson, 2006). The steps of the solution plan must then be held in memory and utilized. Working memory may also contribute to the frequent tendency of children with math

disabilities to undercount or overcount during the math problem solving process (Geary, 1990; Hanich et al., 2001) because the child loses track of where he or she is in the counting process.

Overall, most evidence in this area supports this hypothesis of a relation of working memory and math problem solving, with some evidence for a relation to inhibition as well. While early hypothesis suggested that there might be only a relation of working memory for numerical information to math achievement, researchers have recently found various types of evidence that suggests that both verbal and non-verbal tasks of working memory are related to math performance (Swanson & Alloway, 2012). Passolunghi and Siegel (2001) found that students classified as poor math problem solvers in fourth grade demonstrated deficits in working memory, but no deficits in short-term memory. Multiple correlational design studies have shown that working memory uniquely predicts math problem solving performance in children ages 4 to 5 (Lan, Legare, Ponitz, Li, & Morrison, 2011) and in ages 10 to 12 (Lee, Ng, & Ng, 2009). This has been found to be a substantial relation. LeBlanc and Weber-Russell (1996) found that working memory task performance accounted for 47-52% of variance in math word problem solving for students in first to third grade while Lee and colleagues (2009) found that working memory measures explained approximately 26% of the variation in algebraic problem-solving performance among 11-year-olds.. This relation has been found to remain significant when models include measures of intelligence, reading skills, semantic processing, speed, and inhibition (Cragg & Gilmore, 2014). Lastly, some working memory intervention studies have found that working memory training that significantly improves working memory leads to higher math problem solving at follow up months after the intervention (Holmes, Gathercole, & Dunning, 2009).

Research on the relation of inhibition to problem solving is more mixed than the research examined above. In the research by Lan and colleagues (2011) with 4 to 5 years old, inhibition as well as working memory predicted performance on the math problem solving measure. However, when working memory performance was controlled for, inhibition continued to predict problem solving problems that involved counting but not those involving calculation. In Passolunghi and Siegel's work (2001), those participants with low math problem solving performance also made more intrusion errors on the working memory tasks requiring inhibition of irrelevant information, suggesting that these children with math problem solving difficulties may also have inhibition difficulties. Lastly, Lee and colleagues work (2009) with 10-12 year old children found that inhibition did not significantly predict performance on a math problem solving measure (Lee, Ng, & Ng, 2009). As the authors pointed out, this finding of no relation with inhibition differs from the findings by Passolunghi and Siegel who found a relation of inhibition and math problem solving skills. This may be because Passolunghi and Siegel (2001) included math problem solving items that had information irrelevant to the problem that needed to be ignored while the problems in the Lee and colleagues study did not. Another difference is that participants in the research that did find a relation were older than in the study where no relation was found. Together, these studies indicate that working memory is likely related to math problem solving performance, while inhibition may be important when measures include inhibition of irrelevant information or participants are still mastering the steps of basic math problem solving, such as counting to tell how many items there are.

**Executive functioning and math fluency.** Shifting and inhibition may support successful performance on math fluency tasks, however there is little direct data on the relation of these executive functioning skills and math fluency. Some research exists on math calculation

achievement, which can provide some relevant information but may be different because calculation measures are untimed and involve more complicated calculations compared to timed fluency tasks such as completing as many one or two digit multiplication problems as one can in one minute.

Inhibition may suppress inappropriate strategies during math fluency tasks (e.g., addition when subtraction is required or ineffective counting strategies). Geary, Hamson, and Hoard (2000) asked children to solve addition only by means of retrieval (no counting strategies) and observed that children with math disability (with or without comorbid reading disabilities) committed more errors than the typically achieving control children. The errors were counting-string errors that they attributed to inefficient inhibition of irrelevant associations, such as answering 7 to  $3 + 6$  because it is associated with one of the addends. Geary and colleagues (2000) connected this finding to that of Barrouillet, Fayol, and Lathulière (1997, experiment 2). These researchers presented adolescents with learning disabilities with multiple choice single digit multiplication problems and observed that the adolescents erred more often when the false answers were multiples of one of the operands (e.g.,  $4 \times 6 = 18, 24, 28, \text{ or } 30?$ ) than when they were not (e.g.,  $4 \times 6 = 21, 24, 25, \text{ or } 27?$ ) or when the false answers were not from multiplications tables (e.g.,  $4 \times 6 = 22, 23, 24, \text{ or } 26?$ ). As arithmetic problems may be associated in long-term memory with incorrect answers as well as the correct answers, Barrouillet and colleagues (1997) interpreted the effect they observed as a result of a failure to inhibit competing incorrect answers. Two studies analyzing performance in children in a range of primary grades have found significant correlations between measures of inhibition and performance on a math calculation measure (Gilmore et al., 2013; Lee et al., 2012). However, Censabella, & Noël (2008) found no differences between children with math difficulties and without math difficulties



on any measure of working memory, inhibition, or shifting when students were grouped based on a measure of math calculation and fluency.

Theoretically, shifting ability may also support switching between arithmetic strategies (e.g., addition, subtraction, multiplication) when a task requires the use of multiple arithmetic strategies (Geary, 1990). In line with this, children with arithmetic deficits have been shown to display poorer shifting ability (Bull, Johnston, & Roy, 1999; van der Sluis, de Jong, & Leij, 2004). Therefore, there is some evidence to support the importance of inhibition and shifting for math calculation but conclusions for math fluency in particular are difficult to draw due to lack of research into speeded math calculation performance.

Together, this provides some support for a causal link of improved executive functioning and improved math achievement. For physical activity to have a downstream effect on the outcome of academic achievement through executive functioning changes, however, physical activity would need to affect executive function but these levels of improvement have not been directly linked to improvement in children's academic performance. At this point, the evidence that exists linking executive functions to academic achievement is primarily correlational or relies on quasi-experimental designs and therefore, further research is necessary.

## **Health Indicators and Executive Functioning**

### **Physical Activity and Executive Functioning**

Research into physical activity and cognition suggests that executive functioning may be supported by high levels of physical activity. Early researchers examined if improvements in overall intelligence would be seen with increased PA. This research used standardized IQ tests to

measure general intelligence (Tomprowski et al., 2008). However, the research did not support the idea that PA has a global effect on intelligence. Therefore, a more component based approach began to be taken to assess specific cognitive processes. When examining different cognitive processes, researchers did not find evidence for effects of PA on most specific cognitive processes, including perceptual motor skills, spatial and logical reasoning, and recall of stimuli (Davis et al., 2007; Tuckman & Hinkle, 1986). While some reviews of the literature have also included these studies, due to the definition of executive function described above these studies will not be included in this review. However, research showing improvements on measures of creativity and measures of Planning from the PASS theory was taken to reflect improvements in executive function (Davis et al., 2007; Naglieri & Kaufman, 2001; Tuckman & Hinkle, 1986). Therefore, the primary understanding in the field is that PA influences executive functioning rather than overall intelligence or other specific cognitive processes.

At this time, there have been six randomized control trials in the area of long-term physical activity interventions and executive functioning. In the first study in this area sedentary, overweight children between 7 and 11 years old were randomly assigned to low-dose (20min/day of exercise) high-dose (40min/day of exercise) or control conditions for 15 weeks (Davis et al., 2007). Group differences were found for the Planning scale of the Cognitive Assessment System (CAS; Naglieri & Das, 1997), which is considered by the developer to be an index of overall executive function ability. The high dose group performed significantly better than the control group but no differences were seen between the high and low dose intervention groups. However, this research leaves open the question of whether similar improvements would be seen for children who are not overweight but are not meeting PA guidelines or have low

cardiovascular fitness. It also does not allow for an examination into the effects of physical activity on certain aspects of executive function, however, other studies have begun to do so.

Four randomized control intervention studies have examined the effects of long-term physical activity on inhibition using performance or behavioral measures similar to those used in this study. Chaddock-Heyman and colleagues (2013) randomly assigned 8- to 9-year-old children to receive 60+ min of physical activity, 5 days per week, for 9 months and measured inhibition using performance during a flanker task, where participants are asked to indicate the direction of a target arrow that is sometimes “flanked” by arrows of the same direction and sometimes by arrows of the opposite direction. The PA intervention was the Fitness Improves Thinking in Kids (FIT Kids) curriculum where children completed stations that focused on a specific fitness component (e.g., cardiorespiratory endurance, muscular strength, motor skills) and participate in game play. Chaddock-Heyman and colleagues (2013) did not find significant group and time interaction results for reaction time or accuracy on the flanker task used to measure inhibition. However, the physical activity intervention group showed significant differences from pre to post-test while the control group showed no significant changes. Therefore, the authors argued that this did raise the possibility that the physical activity intervention was responsible for some of the performance improvements across all children from pre-test to post-test. More recent research using the same FIT Kids intervention protocol did find significantly larger improvements in accuracy on a flanker task in the intervention group than the control group (Hillman et al., 2014). Participants in the studies were of similar ages (7-11 years) and the intervention dosages were similar (five days per week for eight to nine months). These two studies leave open the possibility that physical activity interventions may improve inhibition skills in children.

The two other studies differ from the two above in their sample studied. In a study that involved overweight children (8 to 11 years) participants were randomly assigned to either an after-school aerobic exercise program or control group that participated in sedentary activities such as art and board games for five days a week for eight months (Krafft et al., 2014). The intervention group did not show significantly different changes in accuracy or response time on a flanker task as a result of the intervention. Fisher and colleagues (2011) examined the effects of a different model of intervention on inhibition on younger children (5-6 years). They conducted a randomized control trial of a 10-week physical education(PE) intervention that involved two hours per week of aerobically intense PE compared to 2 hours of standard PE (control). Results indicate that the intervention group had higher accuracy on the Attention Network Test, a version of a Flanker task, but that there was no intervention effect on response time. Therefore, the research on changes to inhibition is unclear at this time, with consistent evidence that there is no effect on response time but inconsistent findings on the effects on response accuracy.

Two of the experimental studies have examined working memory performance. In the Scottish study described above (Fisher, 2011), the length of the span remembered was significantly higher in the intervention group and the number of errors on a measure of spatial working memory was significantly reduced in the intervention group. However, the researchers reported that data collected in the same study indicated the spatial working memory measure had low reliability. In the other randomized control trial study in this area, researchers examined the effects of a nine month long aerobic physical activity intervention on working memory performance (Kamijo et al., 2011). Children between the ages of 7 and 9 were randomly assigned to participate in an afterschool physical activity program involving two hours of cardiorespiratory fitness focused activity each school day or a waitlist control group. The youth

completed a modified Sternberg task of working memory, where participants are shown a varying sized string of letters and must remember the string in order to decide if a subsequently presented letter was part of that set or not. It should be noted that there is some debate in the literature about the validity of this task as a working memory task, with some arguing that it instead only taps short term memory (Corbin & Marquer, 2013). The intervention group showed significantly larger increases in overall response accuracy, particularly for the larger sets. This indicates that physical activity may have an effect on working memory skills, although it is not conclusive.

Hillman and colleagues (2014) also analyzed the impact of the 9 month long FITKids intervention described previously on a measure of switching. Children ages 7 to 9 years old were shown items that differed first all on shape and then all on color and asked to decide if another shape was similar to those on shape or color (homogeneous trials). Then the trials involved flexibly switching between making decisions based on shape or color (heterogeneous trials). No difference in improvement was seen between the intervention and control group on participants' response time on either homogeneous or heterogeneous trials. However, the intervention group did improve more than the control group in their accuracy on the heterogeneous trials. In addition, intervention attendance was positively related to the participant's accuracy on the heterogeneous trials. This study indicates a possible effect of chronic physical activity change on shifting but replication is needed.

Overall, there is some evidence for the beneficial impact of physical activity on executive function in youth using both neuropsychological performance and neuroelectrical measures. However, the research into the links between PA and executive function in humans, particularly school-aged youth, is only just emerging. Research examining the effects of real world relevant

long-term physical activity on working memory have sometimes used measures of executive function whose reliability is questionable or whose validity are contested by other researchers. For all of the three areas of executive function to be examined in the current study there are five or fewer studies examining the effects of a long-term physical activity intervention. Therefore, additional research examining the effects of long-term physical activity on all the executive processes using strong experimental design and validated measures is necessary.

These physical activity interventions result in students engaging in more physical activity, which could have an effect on their CVF and/or body composition. However, as was true with the analysis of the literature on physical activity interventions and academic achievement, the studies presented above do not explicitly explore the relation of any changes in cardiovascular fitness and body composition as they relate to the changes in executive function measured. Therefore, a separate analysis of these related health measures is required. Both obesity and CVF can be thought of as attributes that result from similar engagement in healthy behaviors and interactions of genetics and environmental factors. A few investigations on the relation between body fat and executive function have sought to determine the individual and combined contributions of CVF and obesity to executive function (such as Pontifex et al., 2014). This study found no significant interactions between these two factors in the prediction of executive function. In addition, most research has examined these factors and their relation to executive function separately. Therefore, the relations of obesity and CVF with executive function and academic performance will be discussed separately below.

### **Obesity and Executive Function**

The relation between obesity and the executive function of inhibition has not been examined as frequently as the relation with CVF however there is information on the relation to

all three areas of executive function relevant to this study. Eight studies exist in the area of inhibition and obesity in children and adolescents (Reinert, Po'e, & Barkin, 2013). Of these, three utilized measures of inhibition that draw on a similar idea of the construct as the measure use in this study. In a study of inhibition examining the performance 7 to 12-year-old children who completed the Stroop Color-Word task (Buck, Hillman, & Castelli, 2008), higher BMI was not related to scores on the Stroop task. Research using a flanker task did not identify any relation of BMI to accuracy or reaction time on that task (Scudder et al., 2014). However, research with preadolescent children who completed Go and NoGo tasks found that children with higher BMI and children with higher body fat (measured using dual-energy X-ray or DXA) exhibited poorer performance on the NoGo task that required more inhibition (Kamijo et al., 2012). Therefore, the evidence on inhibition is inconsistent, although this may be because all three studies use different measures of inhibition.

The current research into the relation of obesity of shifting have consistent findings, though there is a very small set of research currently. In a study of 24 boys aged 11 to 13, obese children performed worse on the Wisconsin Card Sort task, both overall errors and perseverative errors (Cserjési, Molnár, Luminet, & Lénárd, 2007). There were significant correlations between BMI and perseveration on the Wisconsin Card sort task, which is thought to be a measure of shifting. Similarly, Pontifex and colleagues (2014) found that higher levels of adiposity related to poorer response accuracy for the heterogeneous condition of a shifting task when the variance associated with fitness was controlled. Therefore, there may be an association between obesity and shifting that is beyond the impact of cardiovascular fitness itself. As discussed previously, while cardiovascular fitness and obesity are highly correlated, they are each important aspects of physical health and fitness and therefore are each addressed.

## **Cardiovascular Fitness and Executive Functioning**

It has been well established that a relation between CVF and performance on tasks requiring executive functioning can be found in adult populations (Colcombe & Kramer, 2003; Kramer, Hahn, Cohen, Banich, McAuley, Harrison, Chason, Vakil, Bardell, Boileau & Colcombe, 1999). Researchers have hypothesized and begun to find some evidence that children who are more cardiovascularly fit perform better on attentional tasks that require high levels of executive function. In the past decade, a body of literature examining the relation between CVF and different executive function skills in children has begun to grow.

The majority of studies in this area have examined relations between CVF and measures of inhibition and have found a relation with accuracy and sometimes with response time on the measures. In two studies with higher- and lower-fit children, the group of higher-fit children performed more accurately overall on a flanker task when compared to lower-fit children (Hillman, Buck, Themanson, Pontifex, & Castelli, 2009; Pontifex et al., 2011). In addition, in one of the studies the higher-fit participants were able to maintain a high level of response accuracy across during the more challenging incongruent response trials (Pontifex et al., 2011).. Similar to the results seen in physical activity intervention studies, while there were group differences for accuracy there were no group differences observed for reaction time in either study (Hillman et al., 2009; Pontifex et al., 2011) In addition to this group differences research, Pontifex and colleagues (2014) and Scudder and colleagues (2014) both identified a linear relation between CVF and inhibition, with lower levels of fitness relating to poorer response accuracy for both congruent and incongruent trials of the flanker task. Overall, deficits in performance on inhibition tasks have been observed for lower-fit relative to higher-fit children with both association and group differences methods.



In a study of inhibition using a different measure, researchers examined the performance of 7 to 12-year-old children who completed the Stroop Color-Word task (Buck, Hillman, & Castelli, 2008). Higher levels of aerobic fitness (i.e., the number of laps run on the PACER test) were associated with correctly naming more colors during each Stroop condition, independent of other factors such as age, gender, BMI, or intelligence. These findings also indicate that fitness and inhibition have a significant relationship in preadolescent children.

Researchers have also begun to describe relations between CVF and working memory ability using multiple working memory tasks. Scudder and colleagues (2012) have used the n-back task of working memory. In this task, participants watch as an illustrated black-and-white cow appears pseudorandomly inside one of the six boxes on a screen and then are instructed to respond as quickly and accurately as possible with a right button press if it appeared in the same box as the previous trial during the 1-back condition, and two trials prior for the 2-back condition. In this study, accuracy was significantly greater in higher-fit children for during both the 1- and 2-back conditions. In an intervention study examining similar variables, increases in cardiorespiratory fitness ( $VO_2$ ) resulting from the 9 month physical activity intervention led to improvements in response accuracy on the Sternberg task of working memory, while no significant improvements were seen in the waitlist control group (Kamijo et al., 2011b). Together, these findings reveal evidence that CVF is related to better working memory performance.

In the one study identified that has yet examined the relation between CVF and the executive function of shifting, lower levels of fitness were associated with poorer response accuracy on the more difficult, heterogeneous condition of the switching task (Pontifex et al., 2014). Participants in this study were ages 7 to 10 years of age who were able to perform with at

least 40% accuracy on the measure. No other studies of this question were identified in the literature review search.

Overall, while there are one or two studies indicating a possible relation between working memory and shifting and CVF, research most strongly indicates a relation of CVF and inhibition. However, it should be noted that the only studies found examining working memory and shifting found positive relations. Therefore, there may also be relations to these aspects of executive function. Further research into these aspects of executive function and CVF is therefore needed.

While neuropsychological research provides us information about the link between physical activity and executive functioning as measured using representative tasks, it still provides limited information about how physical activity causes these improvements through physiological changes in the brain. This information has been primarily examined in animal models, and therefore connections to humans must be made carefully. However, it is important to examine what information exists about the potential biological mechanisms for the relation between physical activity and executive functioning.

### **Proposed Biological Mechanisms**

Physical activity has been hypothesized to cause numerous changes in brain neurophysiology that may contribute to changes in cognition and learning outcomes. The majority of experimental research on biological mechanisms for improved learning outcomes and cognition from PA has been done in animal models. Physiological research has focused on cerebral circulation improvements through angiogenesis, changes in neurogenesis and synaptic

plasticity due to changes in concentrations of neurotrophins, and changes in concentrations of neurotransmitters (Trudeau & Shephard, 2010).

As included in the list above, a primary hypothesis for the beneficial effects of PA for cognition and learning is the idea that cognition is enhanced by increased neurogenesis and cell survival in the hippocampus. Neurogenesis is the genesis of new neuron cells. Although most active during pre-natal development, neurogenesis has been shown to continue post-natal in certain areas of the brain of mammals (Altman, 1962). The process of neurogenesis can be modified by environmental factors, such as stress, environmental enrichment, and PA (van Praag, 2008). Cardiovascular PA, such as running, causes an increase in neurogenesis in the dentate gyrus of the hippocampus of rodents, a brain area considered important for learning and memory (van Praag, Kempermann, & Gage, 1999). This research shows a three- to four-fold increase in the production of new neurons in rodents after PA (Kitamura, Mishina, & Sugiyama, 2003; Kobil et al., 2011; Trejo, Carro, & Torres-Aleman, 2001; van der Borght et al., 2006; van Praag et al., 1999). Researchers have found the effect of exercise on the hippocampus to be age-dependent, with more robust hippocampal neurogenesis seen in adolescents and young rats compared to older rodents (Kim et al., 2004). This positive relation between cardiovascular PA and neurogenesis has raised the hypothesis that the new hippocampal neurons may mediate, in part, the improved learning associated with exercise (van Praag, 2008).

In addition to enhancement of neurogenesis, childhood aerobic activity and fitness has been associated with differences in brain volume and changes in the hippocampus (Voss, Nagamatsu, Liu-Ambrose, & Kramer, 2011). During adolescence, the brain continues to develop. Given the dynamic plasticity of the teenage brain, this period may be especially sensitive to the effects of exercise. MRI methods have recently been used to investigate if there

are differences in hippocampal volume in human adolescents based on level of fitness (Chaddock et al., 2010). In this study, higher-fitness children showed greater bilateral hippocampal volumes and superior relational memory task performance compared to lower-fit children. No relation between aerobic fitness, nucleus accumbens volume, and memory was found, which supports the hypothesis of a specific effect of fitness on the hippocampus. In an observational study of 15 to 18 year olds, higher aerobic fitness predicted faster learning on a virtual Morris Water Task and larger hippocampal volumes (Herting & Nagel, 2012). Therefore, this research suggests that differences in hippocampal neurogenesis and resulting hippocampal volume may be related to the effects of exercise on learning and memory.

In relation to this increase in neurogenesis, exercise appears to induce enhanced synaptic plasticity, or the ability of synapses to change the strength of their connection. Long-term potentiation (LTP) is a long-lasting enhancement in signal transmission between two neurons that results from the neurons being stimulated at the same time and is one of several mechanisms underlying synaptic plasticity. Researchers have found that LTP in the hippocampus is influenced by PA (Trudeau & Shephard, 2010; van Praag, 2008). Because memories are thought to be encoded through synaptic plasticity, LTP is the primary physiological model for memory and learning used in experimental research in recent years (Cooke & Bliss, 2006). Research has found that cardiovascular activity enhances this type of synaptic plasticity (O'Callaghan, Ohle, & Kelly, 2007).

Physical-activity mediated enhancement of LTP involves neurotropic factors such as brain-derived neurotropic factor (BDNF) and insulin-like growth factor I (IGF-I). The expression and actions of these neurotropic factors at the synapse between two neurons works to modify synaptic transmission and connectivity (Schinder & Poo, 2000). This modification is an

important requirement for synaptic plasticity. Consistent evidence has been obtained for the role of BDNF in long-term potentiation (Lu & Chow, 1999). Research suggests that PA causes increases in the concentrations of BDNF. For example, several days of wheel running for rodents has been found to increase levels of BDNF in the hippocampus (Neeper, Gómez-Pinilla, Choi, & Cotman, 1996). Research with humans has also found that cardiovascular exercise increases BDNF levels in the bloodstream (Tang, Chu, Hui, Helmeste, & Law, 2008; Zoladz et al., 2008). IGF may also be involved, due to the fact that BDNF gene expression is partially dependent on IGF-1 and because IGF-1 is itself also modulated by exercise (Cotman & Berchtold, 2002). IGF-1 increases in the peripheral nervous system, as shown in humans, and in the brain after exercise (Carro, Nuñez, Busiguina, & Torres-Aleman, 2000; Schwarz, Brasel, Hintz, Mohan, & Cooper, 1996). Overall, PA has been shown to enhance the expression of these neurotropic factors that then enhance LTP.

Because enhancement of neurotropic factors enhances LTP, a physiological mechanism of learning, then enhancing BDNF and/or IGF expression PA can potentially influence learning (Cotman & Berchtold, 2002). Recently researchers have shown that inhibiting BDNF action can block the benefit of exercise on the learning and recall abilities of rodents. While Vaynman and colleagues (2004) found that performance on the water maze test of spatial learning was improved in animals given access to exercise when they blocked BDNF in exercising animals' performance was decreased to the same level as sedentary control animals. It also blocked the effect of exercise on downstream systems regulated by BDNF and important for synaptic plasticity (Vaynman, Ying, & Gomez-Pinilla, 2004). Therefore, evidence that PA can enhance neurotropic factors such as BDNF and IGF is further evidence that PA can have an influence on learning.

A third area of research is the relation of physical activity and increases in systemic blood flow and blood pressure. The hypotheses in this area is that physical activity improves cognition through changes to cerebral circulation. In the short term, this physical activity leads to an increase in cerebral blood flow that is seen in motor areas but also other parts of the cortex (Trudeau & Shephard, 2010). In the long term, regular physical activity may further enhance cerebral blood flow volume, causing an increase in angiogenesis. Angiogenesis is the formation of new blood vessels form from pre-existing vessels. These new blood vessels result in increased blood flow, which may support the increased metabolic demands (oxygen and glucose) of increased brain function. Animal models have shown that exercise stimulates angiogenesis in the hippocampus (van Praag, Shubert, Zhao, & Gage, 2005), cerebellum (Black, Isaacs, Anderson, Alcantara, & Greenough, 1990), and cortex (Ding et al., 2006). However, it should be noted that this hypothesis makes the most sense in relation to improved cognition in older adults who may have decreased cerebral blood flow and may not be relevant to the improvements seen in children who generally have good cerebral circulation (Trudeau & Shephard, 2010). In addition, the animal model research discussed above has been done in aging or older rodents. Therefore, it is unknown whether this biological mechanism could be responsible for any improvements in cognition seen in youth.

Lastly, physical activity appears to have the potential to alter the function of neurotransmitter systems in the brain that may influence executive functioning.

Neurotransmitters are chemicals that transmit messages between neurons by binding to receptors on the receiving neuron and triggering a cascade of changes in that neuron. Animal researchers have identified ways that physical activity enhances the synthesis, metabolism, and availability of different neurotransmitters in the brain. Glutamate is an excitatory neurotransmitter thought to

be involved in learning and memory. In research where rats were exposed to a running wheel, the genes associated with the glutamatergic system were up-regulated (saw an increase in production) while genes related to the gamma-aminobutyric acid (GABA) system were down-regulated (Molteni, Ying, & Gómez-Pinilla, 2002). In particular, the rodents showed an increased in NMDA glutamate receptors. Calcium flow through NMDA receptors is thought to be critical in synaptic plasticity. In addition, the genes that were up-regulated because of this physical activity all have a recognized interaction with BDNF, which again suggests a central role for BDNF in the link between physical activity and brain plasticity. Treadmill running has been found to significantly increase glutamate levels during and for a short while following exercise in rats (Leung et al., 2006). However, levels returned to baseline suggesting this is only an acute effect of physical activity. Secondly, physical activity has also been linked to increased availability of the catecholamine neurotransmitters, dopamine, epinephrine, and norepinephrine (van Praag, 2009). However, again these are immediate neurochemical changes that may be more explanatory of effects of acute exercise on subsequent performance rather than of long-term exercise-induced changes.

Exercise-induced changes in neurogenesis, angiogenesis, and neurotransmitter concentration have all been suggested as possible explanatory mechanisms and have been researched using animal models. Overall research indicates that both changes in neurotrophins and neurotransmitters may lead to changes in synaptic plasticity which are responsible for exercise-induced changes in cognition in youth. These exercise induced changes in cognition may have very important implications for the academic performance of youth. Therefore, it is important to improve our understanding of this area from both a neuropsychological and neuroscience perspective.

## **Purpose of the Present Study**

Despite a common understanding that physical activity is important for a healthy lifestyle, it appears that the majority of youth are not as physically active as recommended by United States public health officials. Adding to the importance for physical health, researchers have begun to collect evidence that physical activity causes physiological changes in the brain that support executive functioning and, separately, that strong executive functioning is beneficial for academic achievement. Although there is a growing literature suggesting that physical activity is positively related to executive functioning and academic performance, conclusions around academic achievement are limited by methodological quality and assessment of mathematics as an overall domain rather than specific skills. In addition, little has been done to directly examine the hypothesis that academic achievement in youth would be improved by physical activity through executive function improvements. The current study aims to add to the field's understanding of the effect of physical activity interventions on academic achievement and executive functioning by using a strong experimental design and measures supported by the educational, neuropsychological, and kinesiology communities and by examining how three executive functioning skills are related to math achievement.

One important area of growth for this literature is improvement in the quality of measures used for both academics, physical activity, and executive functioning. The earliest experimental study was the Trois Rivieres study completed in Quebec (Shepard, 1996). However, the use of student grades for the academic achievement measures leaves some question about the reliability of the data. Many other researchers in this area have also used non-standardized measures of academic achievement (Coe, Pivarnik, Womack, Reeves, & Malina, 2006; Morales et al., 2011; Sigfusdottir et al., 2007). However, when academic achievement is measured based on a non-



standardized measure such as academic grades, the meaning of achievement can vary from site to site and scores can be influenced by expectations of respondents. Similarly, Ahamed and colleagues (2007) used teacher report of the increased physical activity time in their classroom. Measuring the actual increase in physical activity is a strength compared to research that does not take any measures, nevertheless many researchers have done this using self-report or proxy-report measures (Daley & Ryan, 2000; Dwyer et al., 2001; Morales et al., 2011). While report measures are less resource intensive, their use can be problematic because children and adolescents can have difficulty recalling their past activity and adults tend to overestimate children's physical activity levels (Miller, 2004; Sirard & Pate, 2001). Therefore, the use of more objective measures, such as accelerometers (physical activity monitors) in the current study, may be key to ensuring reliable and precise measurements of physical activity.

While some work in the area has used strong standardized measures of academic achievement and physical activity, this work has been more limited in scope than the current study. While Donnelly and colleagues (2009) provided the field with a strong group randomized control trial with strong academic achievement measures and accelerometer data for physical activity, the current study will add to this work by also examining the executive function improvements that were hypothesized but not directly examined. The work by Davis and colleagues (2011) is the only experimental study to examine both executive function and academic achievement and therefore the only study comparable to the scope of this current proposed work. The study was stronger than others in its academic achievement measures, using a standardized measure of academic achievement rather than subjective measures such as student report or teacher assigned grades. However, the participants completed the CAS (Naglieri & Das, 1997) in order to examine executive function which does not allow for examination of specific

executive functioning. Davis and colleagues (2011) also examined this question in an overweight and inactive sample, limiting generalization to inactive but not overweight youth. In addition, many researchers from the kinesiology or neuropsychology fields have measured overall achievement or overall mathematics skills without consideration of the wide diversity of skills in mathematics, much less academic achievement as a whole.

As Etnier and Chang (2009) warn, the fact that many of the measures used in the physical activity literature are not as commonly used in the neuropsychological literature limits the ability for work from different perspectives to be compared. One individual level randomized control trial to examine physical activity's effect on executive function was conducted by Kamijo and colleagues (2011). This study however provides information only on the effect of physical activity on the Sternberg task. This task is not considered one of the twenty nine tests of executive function commonly used by neuropsychologists, as identified by Etnier and Chang (2009) in their review of measurement issues in this body of literature. Therefore, the current study will aim to utilize executive function measures more commonly used in the neuropsychological literature.

The primary goal of the current study was to provide additional information on the potential causal links between both physical activity and executive function and between executive function and academic achievement and the role that these causal links play in the effect of a physical activity intervention by utilizing a rigorous experimental design and measures. The current study is partnered with an existing study by researchers in the fields of Kinesiology and Nursing who are conducting a randomized control trial study involving implementation of after-school physical activity intervention. In this intervention, low-active and low SES female students in 5<sup>th</sup> and 6<sup>th</sup> grade were randomly assigned at the school level to an after-school club

aimed at increasing physical activity and physical fitness. The study examined the following research questions and hypotheses.

### **Research Questions and Hypotheses**

1. Do students in the intervention group have higher executive functioning skills than the control group as the result of an after-school physical activity intervention?

*Hypothesis 1:* Students in the intervention group will have significantly more improved performance on the tasks of inhibition and working memory than the control group. Performance on the measure of shifting will not be different for the groups.

*Rationale:* According to the executive function hypothesis, aerobic exercise will have a differential improvement on executive functioning over other cognitive processes. Research with school aged children has found differences on an overall measure of executive functioning, the CAS Planning index (Davis et al., 2011). While the current population will not be recruited for overweight status, they will be recruited based on low current physical activity. Three of the four PA intervention studies conducted with school aged children have identified possible intervention effects on accuracy on the flanker task of inhibition used in this study (Chaddock-Heyman et al., 2013; Fisher et al., 2011a; Hillman et al., 2014). Two studies have identified improved performance on measures involving working memory (Sternberg Task and CANTAB battery) in children following PA interventions (Fisher et al., 2011b; Kamijo et al., 2011a). While recent research by Hillman and colleagues (2014) has shown differences in accuracy on a measure of shifting, this research has found evidence for differential changes only on heterogeneous trials and the measure used will not differentiate between heterogeneous and

homogeneous trials and response time and accuracy but instead provides a single score. In addition, because preliminary research suggests that children may reach mature shifting ability by 10 or 11 years of age (Anderson, 2002; Huizinga, Dolan, & van der Molen, 2006) and participants in the current study will be ages 10 to 15, a shifting effect is not expected to be found.

2. Do students in the intervention group have higher math performance than the control group as the result of an after-school physical activity intervention?

*Hypothesis 2:* Significant differences between groups will be seen in both academic measures assessed (math fluency and applied problems), with the intervention group showing more improvements than the control group.

*Rationale:* The proposed mechanism of change in this study is that changes in physical activity will result in changes in executive functioning (executive function hypothesis) and that this will also show in differences in academic achievement because of the link between executive functioning and academic achievement. Executive functioning is theorized to be important for behaviors that lead to success in the classroom as well as possibly influencing children's ability to apply knowledge and maintain attention during an academic test. Multiple correlational studies have found a link between level of physical activity and academic achievement (Dwyer, Sallis, Blizzard, Lazarus, & Dean, 2001; Fox, Barr-Anderson, Neumark-Sztainer, & Wall, 2010; Kantomaa, Tammelin, Demakakos, Ebeling, & Taanila, 2010; Nelson & Gordon-Larsen, 2006; Sigfusdottir, Kristjansson, & Allegrante, 2007). The afterschool PA intervention work by Davis and colleagues (2007) which showed intervention effects on overall math achievement was completed with an overweight BMI as criteria for inclusion, which was not the case for this

study. However, participants were recruited based on low current PA levels which means they are likely to also have higher body fat than on average. Relatedly, researchers have found a relationship between BMI and math achievement (Kamijo et al., 2012). The participants are also likely to have lower cardiovascular fitness due to being selected for low PA levels and differences in math achievement by levels of cardiovascular fitness have been seen (Castelli, 2007). If the PA intervention does cause changes in body composition and/or cardiovascular fitness, then this change may be seen here in math achievement.

3. What is the relation between change in executive functions and change in math achievement? If there are differential improvements in math achievement by intervention group, are these differences mediated by a change in executive functioning?

*Hypothesis 3a.* Improvements in math problem solving will be mediated by improvements in working memory and inhibition.

*Hypothesis 3b.* Improvements in math fluency will be mediated by improvements in shifting and inhibition.

*Rationale:* Performance on tasks measuring all three executive functions measured in this study has been linked to mathematical performance overall (Bull & Scerif, 2001; St Clair-Thompson & Gathercole, 2006) However, while executive functioning is thought to be distinct yet unified, academic tasks are complex and engage multiple skills. Therefore, it is possible to hypothesize distinct connections between executive functioning and certain math skills. As discussed previously, associations of working memory and inhibition with math problem solving and of inhibition and shifting

with math calculation have been found (Geary, Hamson, & Hoard, 2000; Lan et al., 2011; LeBlanc & Weber-Russell, 1996; Lee et al., 2009; Passolunghi & Siegel, 2001; Swanson & Alloway, 2012). The evidence suggests that inhibition will be related to math problem solving when irrelevant information is included in the math problems as is the case for the measure used in this study. Though the findings for math calculation may differ from how executive functions relate to speeded single digit math fluency, theory suggests that if improvements in switching are seen than improvements on the measure of math fluency will also be seen, as the timed and mixed problem nature of the measure (random addition, subtraction, and multiplication) requires participants to quickly switch math computations.

## **CHAPTER 3: METHODS**

### **Study Approval and Funding**

The study was approved by the Institutional Review Board of Michigan State University, IRB # 14-258. School administrators also provided approval for the study to be conducted in their respective districts and signed Memorandums of Understanding that outline the agreement for both the Girls on the Move study and this study. Funding for the study came from the American Psychological Foundation's David H. and Beverly A. Barlow Grant.

### **Trial Design**

The Girls on the Move (GOTM) project is a group randomized trial study. Schools were included in the GOTM study based on the following criteria: (1) location in an urban community setting; (2) enrollment size in any combination of 5<sup>th</sup>, 6<sup>th</sup>, 7<sup>th</sup>, and/or 8<sup>th</sup>-grade that will allow for around 65 students to be enrolled in the physical activity club in each school; and (3) a student body comprising at least 50% minority versus non-minority race or ethnicity and a similar percentage enrolled in the free and reduced lunch program. From the schools identified as meeting these characteristics, schools were matched based on the following criteria: school size, percentage of students who are white vs. non-white, and proportion of students receiving free and reduced lunch. Then, the schools in the pairs were randomly assigned to the intervention or control group by the GOTM team. Data was collected at baseline prior to the intervention start and again at the conclusion of the intervention, 17 weeks later.

## Study Procedures

**Recruitment Procedures.** Members of the GOTM team scheduled a time to hold an assembly at each school to provide information about the study and invite female students to participate. A brief recruitment video was played to highlight reasons for participating. Students were also informed that their school would be randomly assigned to continue with school offerings or to receive the after-school physical activity club. They were told that all girls, however, could receive incentives for participating in the data collection activities during two times in the year. Following this informational session, the team members distributed packets containing consent/assent forms for the GOTM study and this study and a screening tool. They were asked to return the completed forms in the next day to two and received \$5 for returning the forms, regardless of whether they chose to participate. Participation was then determined on a first-come, first-served basis, as there was limited space in the physical activity club.

Power analysis was conducted to determine the necessary subsample of the total GOTM participants for the completion of this study. A MANCOVA with two levels and two dependent variables, for use in analysis of the two math achievement measures, was conducted in G-POWER to determine a sufficient sample size using an alpha of 0.05, a power of 0.80, and a small effect size ( $f^2 = 0.10$ ) (Faul et al., 2008). Based on the aforementioned assumptions, the desired sample size was 100. Power analysis for a MANCOVA with two levels and three dependent variables, for use in analysis of the three executive function measures, was conducted in G-POWER to determine a sufficient sample size using an alpha of 0.05, a power of 0.80, and a small effect size ( $f^2 = 0.10$ ). Based on the aforementioned assumptions, the desired sample size was 114. Each GOTM club aims for 65 students to enroll in the club. One control and intervention school would therefore generate a sample of 130 students. However, not all students



who chose to enroll in the GOTM study would necessarily also chose to participate in this study. In addition, attrition due to participants changing schools or no longer desiring to participate was anticipated. Therefore, participants were recruited from four schools (two intervention and two control) to ensure a sample sufficient to conduct analysis for the first two research questions.

**Intervention Procedures.** Participants in the intervention groups received a comprehensive 17-week school-based intervention aimed at increasing participants' minutes of MVPA. The intervention group participated in an after-school PA club offered three days per week for 90 minutes each day. The club offers enjoyable PA designed to provide the girls with sixty minutes of MVPA and to help them improve PA skills. Two thirty minutes intervals designed to generate MVPA were provided each session. The various types of MVPA offered each day included choice of: 1) fun games (e.g. tag, kickball, boot camp stations, hula hoops, jump ropes, and capture the flag) 2) dance/aerobics (e.g. dance video games projected on a large screen, Zumba and other dance fitness routines, line dances popular among girls, aerobics, and Pilates); and 3) walking or sport skills (e.g. soccer, basketball, volleyball, lacrosse, running, tennis, martial arts, track, floor hockey, badminton; ultimate disc). In addition to the sixty minutes of MVPA, the students engaged in five minutes of warm up and stretching and were provided a healthy snack.

The club instructors also encouraged the girls to supplement MVPA on their own outside of the club and to think of which activities they completed in the club that they would enjoy outside of the club. In the control condition, students did not receive any after-school programming other than the programming currently offered by their school and community.

Students in the intervention group also received two face-to-face and individualized motivational counseling sessions with the school nurse that lasted for twenty minutes each. The first session occurred at the beginning of the intervention period and the second occurred during

the final month of the intervention. These motivational interviewing sessions were based on Self-Determination Theory (SDT). The three basic needs hypothesized by SDT to drive behavior (competence, autonomy, and relatedness) were addressed during these sessions (Markland & Vansteenkiste, 2007). The motivational interviewing focused on motivation for exercise and attendance at the program and does not address motivation for academic performance. The nurses received two eight-hour days of training in motivational interviewing conducted by a trainer from the Motivational Interviewing Network of Trainers. They received training on the theory, reflective listening, and adolescent development and participate in motivational interviewing demonstrations. Intervention participants also received tailored motivational messages via a computer program at the halfway point of the intervention. Control groups continue to receive their usual school offerings and do not receive any motivational counseling regarding motivation for exercise.

**Treatment Fidelity Procedures.** Process evaluators evaluated the fidelity of the physical activity club using an observation tool and quantitative checklist. Two process evaluators evaluated the club on different days of the same week. These evaluations occurred across the intervention period, separated by approximately one month, for a total of 10 fidelity observations. The intervention coordinator intervened with the club instructors as necessary based on the observations to promote fidelity. In addition, every other week five students were randomly asked to wear accelerometers (for MVPA data collection) during the club to provide a direct measure of how much physical activity students engaged in during the club periods. In addition, to ensure student attendance at the club, girls who missed a full week or more than an average of two club sessions in a week during the first three weeks received a phone call to discuss with parents possible ways to help the girls attend the club more regularly.

**Data Collection Procedures.** Demographic information about student age, grade, and ethnicity was collected in the screening form sent home and returned with the consent forms. All outcome measures were assessed before and after the intervention. Baseline and intervention data for all measures were completed within the same three-week time window. After baseline data had been collected, guardians were contacted by phone to inform them of their school's randomization status. Please see the table in Appendix C for more details regarding the overall schedule of data collection.

For the academic and executive function measures of this study, students were assessed during the school day, in the facilities available at each of the school buildings. They were assessed on the measures of academic performance and executive function one-on-one with research assistants, however due to space limitations in the school the participants were assessed in the same room as one to three other participants. The intervention and assessment teams of the GOTM study functioned independently, so that assessment of body mass index, body fat percentage, cardiovascular fitness, and MVPA is completed by blind assessors. Assessment of academic achievement and executive functioning was completed primarily by research assistants blind to the intervention assignments status of participants. However, due to the need to complete data collection quickly and in collaboration with the school schedule, some assessments were completed by the graduate student researcher who was aware of the intervention status.

The participants were assessed using the standardized assessment procedures outlined by the developers of each measure. The data collection was completed by undergraduate and graduate research assistants. They completed training on the academic measures with the graduate student researcher and completed the NIH Toolbox trainings online. The research

assistants were cleared for data collection after successfully administering the measures to the graduate student researcher using standardized administration protocols.

The GOTM measurement coordinator trained and supervised the GOTM measurement staff. The staff were certified for data collection only after successful practice demonstration. Either the measurement coordinator or project manager was present during data collection sessions to ensure fidelity to the measurement protocol. The same group of trained staff collected data for each matched pair of intervention and control schools. Information was collected in waves, with only half of the data being collected from half of the participants at one time.

### **Participants**

Participants are female students in the Kentwood and Lansing Public Schools in Michigan during the 2014-2015 school year. Two Kentwood Middle Schools (5<sup>th</sup> through 7<sup>th</sup> grade) were matched and two Lansing K-8 buildings were matched. The two Lansing buildings are themed academies. Lansing STEM academy is a school focused on science, technology, engineering, and math. Lansing Gardner academy is focused on leadership, law, and government. Participants all met the following criteria: identified as female, were in 5<sup>th</sup> through 8<sup>th</sup> grade, could participate in the PA club three days a week for seventeen weeks, expected to be available for the 9 month follow up, and reported that they could read and speak English. Girls were not eligible for the GOTM study if they are already involved in organized PA three or more days a week or had a health condition that precludes safe engagement in physical activity. There were no additional inclusion or exclusion criteria for participation in this study beyond the GOTM criteria.

## Measures

### Math Achievement

Academic achievement was assessed using norm-referenced test scores on measures of math fluency and math problem solving. Math fluency and problem solving skills were measured using subtests from the norm-referenced, individually administered Woodcock-Johnson III Tests of Achievement (WJ III; Woodcock, McGrew, & Mather, 2001). The Math Fluency (WJMF) subtest requires students to perform simple one digit and two-digit addition, subtraction, and multiplication calculations under timed conditions. The Applied Problems (WJAP) subtests, used as a measure of math problem solving, requires students to solve orally presented word problems with the aid of pencil and paper. Content areas on the WJAP measure include counting, money, time, and geometry. Form A was used at pretest and Form B at posttest, to minimize practice effects.

The WJ III was normed on a large sample matched to the demographics of the United States in the 1990s in terms of gender, race, geographic region, and type of school (McGrew & Woodcock, 2001). It is therefore considered adequately representative of the North American school-aged population (Cizek, 2003). All reported reliabilities range from .80 and .90s for individual tests. Median test-retest reliabilities ranged from .70 to the .90s. For the speeded math fluency subtest, reliability estimates were calculated using a Rasch approach. For the applied problems subtest, split-half reliabilities were calculated. In addition, information provided by the test developers suggests that the two forms (A and B) are equivalent enough to be used interchangeably.

Scoring of applied problems items as correct or incorrect was completed by the research assistants during testing to determine basal and ceiling levels. Raw scores were totaled by the research assistants and confirmed by the graduate student researchers. Raw scores were converted into grade based equivalents by the graduate student researcher using the Compuscore computer program. Grade based equivalents were used to ensure that any participants who had repeated a grade were not being compared to students of the same age who had exposure to higher grade level material.

### **School Attendance**

Student school attendance records for the dates during with the intervention occurred were collected from school secretaries with access to the school's attendance records system. Using this data, the number of school days each student was present during the intervention period was calculated. To standardize attendance information across schools and grades, participants were marked as having one absence if they were marked absent for the day or absent for more than half of the periods of the day. Participants who were marked absent for the AM or PM or for half of the periods of the day (e.g. absent periods 4, 5, and 6 of a 6 period school day) were marked as having one half of an absence. No difference was made between excused and unexcused absences. Absences for field trips or school related purposes were not marked as absences. These data were collected to provide information about the level of education opportunity, because attendance at school is required to benefit from instruction.

### **Executive Functioning**

Executive functioning (EF) is a family of mental processes thought to be used to concentrate on to achieve a goal. Although there is some debate about the organization of these

processes, in general executive function is thought to include (response) inhibition/inhibition, working memory, and cognitive flexibility/set shifting/updating (Diamond, 2013). The measures chosen for this study are thought to require and measure these aspects of EF (Weintraub et al., 2013), however, it should be noted that there is debate in the field about what aspects of EF are required and assessed by different tasks. Measures from the National Institute of Health (NIH) Toolbox for the assessment of Neurological and Behavioral Function were used to assess executive function (Gershon et al., 2013). The NIH Toolbox provides royalty-free, brief assessment tools that were chosen for inclusion based on applicability across the lifespan, psychometric strength, brevity, ease of use, and applicability in diverse settings (Weintraub, Sandra et al., 2013). These measures were subjected to an iterative process of measure development that included pilot testing to modify existing measures to satisfy the criteria of the Pediatric, Accessibility, and Cultural Working Groups involved in the production of the NIH Toolbox measures. The measures were validated in English using a national standardization sample (n=476) that was representative of the U.S. population based on gender, ethnicity, race, and socioeconomic status. Measures had to meet the following criteria to be included in the Toolbox: minimum convergent validity of .05, divergent validity that is 0.1 less than with the other tests, and split-half internal consistency reliability of 0.65. Test-retest reliability for the three measures to be used in this study was found to be strong (ICC=.89-.96). Psychometric information on the specific NIH Toolbox measures used in this study can be found in the following sections on each measure.

**Inhibition.** To assess inhibition, a three-minute Flanker Inhibitory Control and Attention Test from the NIH Toolbox was used (NIH FIT). This is a version of the Eriksen flanker task that is frequently used (Eriksen & Eriksen, 1974). For this task, the participant must focus on a

center stimulus arrow in a line of arrows and press the keyboard arrow keys to indicate the direction the center arrow pointing. Sometimes the middle stimulus is congruent (pointing in the same direction as the “flankers” or homogeneous trial) and sometimes it is incongruent (pointing in the opposite direction or a heterogeneous trial). Therefore, the participant must selectively focus on the center arrow and inhibit attention and any responses to arrows flanking it. Twenty trials were conducted with each participant, which took about three minutes.

Scoring was completed by the NIH Toolbox program as outlined in the NIH Toolbox Scoring and Interpretation manual. A 2-vector scoring method that uses both accuracy and reaction time is used by the NIH Toolbox system, where each of these “vectors” ranges in value between 0 and 5. Because reaction time data tends to have a positively skewed distribution, a log (Base 10) transformation is applied to each participant’s median reaction time score to create a more normal distribution of scores. The computed score combines the two vector scores and is converted to a standardized score with mean of 100 and standard deviation of 15 based on the normative data for participants of the same age.

Development testing found that scores on this measure were significantly positively related to scores on the D-KEFS Inhibition measure in 8–15 year olds,  $r(81) = 0.34$ ,  $p = 0.002$ . This is positive evidence of convergent validity to other measures of inhibition. Intraclass correlations for this measure show strong test-retest reliability ( $ICC = .92$ ). However, validation research did show a practice effect ( $mean = .51$ ,  $SD = 1.14$ ; Zelazo et al, 2013).

**Working Memory.** The construct of working memory refers to the ability to hold and manipulate information and then hold the products of that manipulation. To measure working memory, the List Sorting Working Memory Test from the NIH Toolbox was used (NIH LS). In this task, participants were presented with a list of stimuli both visually (picture) and auditory



(recording of a one-word description of the stimulus) on a computer monitor. It is important to note that this is different from many working memory tasks, which use either visual or auditory stimuli. The stimuli were presented one at a time at a rate of two seconds per stimulus.

Participants were asked to say them back in size from smallest to largest. First this is done within a single category (e.g. either animals or foods), called the 1-list phase. After the 1-List phase, all participants proceed to the second phase of the test (the 2-List phase), where they see lists of items drawn from both categories (i.e., food and animals). Participants are instructed to reorder and repeat the stimuli first from one category, then the other, in order of size within each category. Lists in the both phases increase from a 2-item list to a 7-item list, for a total of six levels of list length.

For both phases, a second trial at a given list length is administered when participants fail the first trial for each list length. Several discontinue rules are embedded in the software. If a participant does not complete a practice item correctly, the test itself is not administered. If the participant has a score of 0 points as the sum of the two sets of 2-stimulus items (pig-mouse and bird-cow; banana-watermelon and apple-blueberry) or has a score of 1 point as the sum of the two sets of 2-stimulus items (e.g. pig-mouse and bird-cow; banana-watermelon and apple-blueberry) and a score of 0 points on the set of 3-stimulus items (pumpkin-strawberry-banana and dog-horse-rabbit) in the 1- list condition, the test is discontinued and the 2-list condition is not administered. The test takes approximately 10 min to administer.

As described in the NIH Toolbox Scoring and Interpretation manual, test scores consist of combined total trials correct on the 1-List and 2-List phases of the task. Participants receive a score of 2 points if they are correct on the first trial and a score of 1 point if they fail the first trial at that length but pass the second trial. The total number of items correctly recalled and

sequences were summed by the NIH Toolbox program and then converted to the nationally normed, age based standardized scores.

This task has not been found to have any significant practice effect over an average 2-week test–retest interval for children ages three to fifteen, mean practice effect = 0.15, SD = 1.68,  $t(65) = .72$ ,  $p = .48$ . (Tulsky et al., 2013). Correlations with the WISC-IV Letter-Number Sequence measure were significant ( $r = .57$ ) for children ages 8 to 15, an indication of the concurrent validity with similar accepted working memory tasks. The correlations with the discriminant validity measure (the PPVT-4/receptive vocabulary) were also significant however ( $r = .45$ ) indicating an overlap with this construct.

**Shifting.** The NIH Toolbox Dimensional Change Card Sort Test (NIH DCCS) was used to assess participant’s ability to shift or update a set after a rule change is made. This task was designed by Zelazo and colleagues (Frye, Zelazo, & Palfai, 1995; Zelazo, 2006) and based on Luria’s work on rule use. Intraclass correlations calculated for this measure during development show strong test-retest reliability ( $ICC = .92$ ). In addition, the NIH DCCS showed no practice effect over an average two-week test–retest interval, mean practice effect = 0.04, SD = 1.20,  $t(47) = 0.21$ ,  $p = 0.83$  (Zelazo et al., 2013).

In this task, participants were presented with two target pictures that varied on two dimensions. The target visual stimulus must be matched to one of two choice stimuli according to shape or color. This version of the sorting task consists of four blocks: practice, pre-switch, post-switch, and mixed. Participants were required to get three out of four practice trials correct in order to move to the next blocks and if they failed to meet this the first time, the four practice trials were repeated up to three times. Once they met this criterion, they received a comparable series of practice trials for the other dimension (color, in this example). Children who met

criterion for each dimension proceeded to the test trials, which were similar in structure but involved different shapes and colors and no feedback. During the pre-switch block, they were asked to match a series of pictures to the target picture according to one dimension. Then, for the post-switch block they are asked to match according to the other dimension. Those who succeed with this switch also receive a mixed block, in which color was the matching criteria on the majority of trials with occasional, unpredictable shifts to shape. The relevant criterion word (color or shape) appeared on the screen for all participants and was also delivered orally via the computer speaker. The test took approximately four minutes to complete.

Performance was scored on the test trials that each participant received, whether these included only the pre-switch block, both pre- and post-switch blocks, or all blocks. For children who received only the pre-switch block, or only both pre- and post-switch blocks, the final score was equal to the accuracy score. For children who received all blocks, and whose accuracy across all trials was less than 80%, the final score was equal to the accuracy score. Accuracy was scored on a scale from 0 to 5. Children were given 0.125 points (5 points divided by 40 total task trials: 5 pre-, 5 post-, and 30 mixed-block trials) for every correct response they made on trials they received. In other words: Accuracy Score =  $0.125 * \text{Number Correct Responses}$ . For participants who received all blocks and whose accuracy was 80% or higher, an RT score was also calculated based on each participant's median RT on correct infrequent trials during the mixed block. On tasks like the DCCS, older children and adults typically slow down (increase in RT) in order to respond accurately (Zelazo, 2006). Therefore, a two-vector method is used that incorporates both accuracy and, for participants who maintained a high level of accuracy (> 80% correct), reaction time (RT). Like the accuracy score, the RT score ranged from 0 to 5. Trials with RTs lower than 100 ms or greater than 3 SDs from each participant's mean RT were

discarded as outliers prior to calculation of median RT. Because RT tends to have a positively skewed distribution, a log (Base 10) transformation was applied to each participant's median RT score, creating a more normal distribution of scores. Based on the validation data, the minimum RT for scoring was set to 500 ms and the maximum to 3,000 ms. Median RTs that fell outside of this range but within the allowable range of 100-10,000 ms were truncated for the purposes of RT score calculation. Scoring of the validation data had indicated that this truncation did not introduce any ceiling or floor effects. Log values were algebraically rescaled from a  $\log(500) - \log(3,000)$  range to the 0-5 range. Note that the rescaled scores were reversed such that smaller RT log values are at the upper end of the 0-5 range whereas larger RT log values are at the lower end. Once the rescaled RT scores were obtained, they were added to the accuracy scores for participants who achieved the accuracy criterion of 80% or better. All participant scores were converted to a standardized score based on their age, with a mean of 100 and a standard deviation of 15.

### **Intervention Dosage**

The amount of intervention, or dosage, received by each participant will be measured through their attendance at each session. Club coaches were responsible for taking attendance each session. The number of total sessions provided at the school was also recorded. The number of sessions attended was divided by the total number of sessions provided at that school to determine a club attendance percentage for each participant.

### **Physical Activity**

The number of minutes of MVPA per hour engaged in by both intervention and control subjects was assessed using a small accelerometer (ActiGraph GT3X-plus). Accelerometry-

based motion sensors have become one of the most commonly used methods for assessing physical activity. Accelerometers provide quantitative information about the vertical accelerations of the trunk or other body segments at user-specified time intervals. Therefore, they can be used to evaluate the frequency, intensity, and duration of physical activity over specified time intervals. A large number of studies investigating the validity of accelerometers by comparing the data to indirect calorimetry, doubly labeled water, and direct observation (Troost, 2007) have reported strong positive correlations between accelerometer output and energy expenditure. It should however be noted that a limitation of accelerometers is that they cannot account for the increased energy expenditure from activities such as walking upstairs or an incline, cycling, weight lifting, or carrying objects. Because of this, they can underestimate physical activity and energy expenditure. Regardless, accelerometers have significant advantages over recall and self-report based methods and are considered a strong tool for measuring physical activity in free-living children and adolescents.

The accelerometer used in this study is the Actigraph (Actigraph LLC, Fort Walton Beach, Fla). The Actigraph has documented evidence of validity and inter-instrument reliability in children and adolescents. Concurrent validity studies using heart rate monitoring, direct observation, indirect calorimetry, whole-room calorimetry, and doubly labeled water have all demonstrated the Actigraph to be a valid measure of physical activity in children and adolescents (Troost, 2007). In addition, studies have shown the Actigraph to have acceptable to high levels of intraunit and interunit reliability (Troost, McIver, & Pate, 2005).

To collect the physical activity monitor data, a staff member provided instructions on wearing the accelerometer with each girl and played a two-minute video to reiterate the importance of wearing the monitor. Girls were asked to wear the monitor attached to an elastic

belt on their right hip from getting out of bed in morning until returning to bed to sleep at night for seven consecutive days, but not when bathing or swimming. Among adults, three to five days of monitoring is required to reliably estimate habitual physical activity. However, among children and adolescents, the number of monitoring days required has been found to be between four and nine days. To remind them, participants received an automated phone call to their homes every morning before school and at 11 A.M. on Saturday and Sunday until the monitor was returned. Monitors were set to begin data collection at 5:00 A.M. on the day after they are distributed to participants at school.

Minutes of MVPA per hour in a week was estimated from the acceleration counts recorded divided by the total wear time during that week. The activity counts created by the monitors provide an estimate of the intensity of measured activities during each time period. Analysis was conducted using count cut-points (15-second epochs) as shorter time-sampling intervals better capture changing movement and have the best sensitivity and specificity. According to the protocol for the GOTM Intervention, count thresholds where 574-1002 counts in a 15 second period are moderate intensity and >1003 counts/15 seconds are vigorous PA were used. The effect of the physical activity club intervention on the student's overall physical activity level was assessed by summing the MVPA counts from 6 A.M. to midnight to determine overall minutes of MVPA and dividing this by the time the monitor was worn. Data were excluded if the monitor is taken off during waking hours, as determined by 60 or more consecutive minutes of continuous zeroes (Robbins et al., 2013). As discussed previously, to measure the level of PA specifically during the club, five girls per school were randomly selected every other week to wear the accelerometer during a meeting of the club. Both overall physical activity and physical activity during club meetings were determined using this measure.

## **Body Composition**

Body composition refers to the relative amounts of the components that make up a person's total body weight. The major contributors to body weight are the fluid, muscle, bone, and fat content. In most body composition assessment models, all lean components (e.g., fluid, muscle and bone) are combined into the fat-free mass (FFM). While a positive relationship between physical activity and body composition is established in adults, the relationship is surprisingly less clear in children. However, some researchers examining the relations between activity measures and measures of body composition have identified a relationship (Rennie et al., 2005). Because body composition may have a relation to the physical activity intervention and to cognitive functioning and academic performance, this information was a potential covariate for analysis. Selection of specific body composition measures was determined by the GOTM intervention based on the research questions and considerations of feasibility in a school setting. For this study, body mass index (BMI) and percent body fat will be used.

**Body Mass Index.** BMI is an indicator of weight-for-height. Height without shoes was measured to the nearest 0.1cm and weight was measured to the nearest 0.1kg (Robbins et al., 2013). Both measurements were completed behind a privacy screen. BMI raw scores were calculated from these measures (weight in kg/height in meters squared). From this, z-scores for BMI were determined by the GOTM team using a SAS program for Center for Disease Control Growth Charts (available from the National Center for Chronic Disease Prevention and Health Promotion). This standardized z-score was used for the data analysis. BMI is used frequently in public health research to monitor overweight and obese status, however, BMI is thought to estimate fatness with an error of >5% (Plowman & Meredith, 2013).

**Body Fat Percentage.** Body fat percentage was measured behind a privacy screen. It was measured to the nearest 0.1% using a foot-to-foot bioelectric impedance analysis (BIA) scale. These BIA scales measure the impedance or resistance to a small electrical current as it travels throughout the body's water-containing tissues. This measure of resistance to the current was used to calculate total body water. Because most of the body's water is in fat-free mass, this generates an estimate of fat-free mass. The advantages of this method include ease of use and relatively low cost. One limitation of this method is that the algorithms for conversion do not account for changes in composition during childhood and adolescence and for racial/ethnic variations in body composition (Institute of Medicine et al., 2012). Bioelectric impedance analysis methods have been found to have an error of around 3% (Wells, 2005).

### **Cardiovascular Fitness**

Cardiovascular fitness (CVF) is the ability to perform whole body exercise at moderate to high-intensity for an extended period (Institute of Medicine, 2012).  $VO_2$  max is considered to be a measure of cardiovascular fitness. It is a measure of the maximum capacity of an individual to transport and use oxygen during exercise. The Progressive Aerobic Cardiovascular Endurance Run (PACER) is a variation on the shuttle run commonly used as measures of cardiovascular endurance and aerobic capacity. The PACER was used to determine  $VO_2$  max for the purposes of this study. In this measure, lines are placed either 15 or 20 meters apart based on space availability and participants complete the task in groups of up to six students. The participants run repeatedly from one line to another when it is indicated by audio cues. Each time they complete a run in one direction it is considered a lap. The time between the audio cues decreases periodically while the distance remains the same. Eventually, the participant is unable to reach the other line before the time is up. When a participant has not completed two laps within the



allotted time then they are considered finished with the PACER and the number of successful laps is recorded. The number of laps completed was converted to estimated VO<sub>2</sub> max by the GOTM team for analysis, with the higher number of laps meaning greater cardiovascular fitness.

### Demographic Information

Information about participant age, grade, participation in free and reduced lunch, and reported ethnicity was collected from the screening form by the GOT team.

Table 1  
List of Constructs and Measures to be Used in Analysis of Research Questions

<b>Area</b>	<b>Construct</b>	<b>Measure</b>
<b>Math Achievement</b>	Math Fluency	Woodcock-Johnson III Test of Achievement-Math Fluency (WJMF)
	Math Problem Solving	Woodcock-Johnson III Test of Achievement-Applied Problems (WJAP)
<b>Educational Opportunity</b>	School Attendance	Number of days absent
<b>Executive Function</b>	Shifting	NIH Toolbox- Dimensional Change Card Sort
	Working Memory	NIH Toolbox- List Sorting
	Inhibition	NIH Toolbox- Flanker Inhibitory Test
<b>Physical Activity</b>	Moderate to Vigorous Physical Activity	ActiGraph GT3X-plus accelerometer
<b>Body Composition</b>	BMI Z-score	Calculated from Height and Weight and transformed to z-score
	Body Fat Percentage	Bioelectric impedance analysis (BIA) scale
<b>Cardiovascular Fitness</b>	VO <sub>2</sub> max	PACER
<b>Intervention Dosage</b>	Club Attendance Percentage	Number of club sessions attended divided by number of sessions provided

Table 2  
 Research Questions, Statistical Questions, Analysis, and Variables

<b>Research Question</b>	<b>Statistical Question</b>	<b>Analysis</b>	<b>Variables</b>
1. Do students in the intervention group have higher executive functioning skills than the control group as the result of an after-school physical activity intervention?	Does the mean of the executive functioning measures change over time differently between the intervention and control group?	Repeated Measures ANOVA	Pre and Post-test of NIH-LS, FIT, and DCCS
2. Do students in the intervention group have higher <u>math performance</u> than the control group as the result of an after-school physical activity intervention?	After controlling for pre-test math achievement, is there a statistically significant difference on post-test math achievement by intervention group?	Multiple Analysis of Covariance (MANCOVA)	Pre and Post-test of WJMF and WJAP
	Does the mean of the math achievement measures change over time differently between the intervention and control group?	Repeated Measures ANOVA	Post-test of WJMF or WJAP  Covariate: Pre-Test WJMF or WJAP
3. What is the relation between change in executive functions and change in math achievement? If there are differential improvements in math achievement by intervention group, are these differences mediated by improvements in executive functioning?	Are there significant indirect effects of executive function measures in a model of the effect of intervention on the math measures?	Simultaneous multiple mediator model	Intervention Status  Change scores of WJMF or WJAP Change scores of NIH LS, FIT, and DCCS

## CHAPTER 4: RESULTS

### Sample Characteristics

The original number of participants recruited was 167. These were a subsample of the GOTM study sample. A total of 17 participants were removed from the sample used for data analysis. Information about losses after randomization can be found in Figure 4 below. The sample utilized for analysis was 150 participants with 78 control and 72 intervention participants.

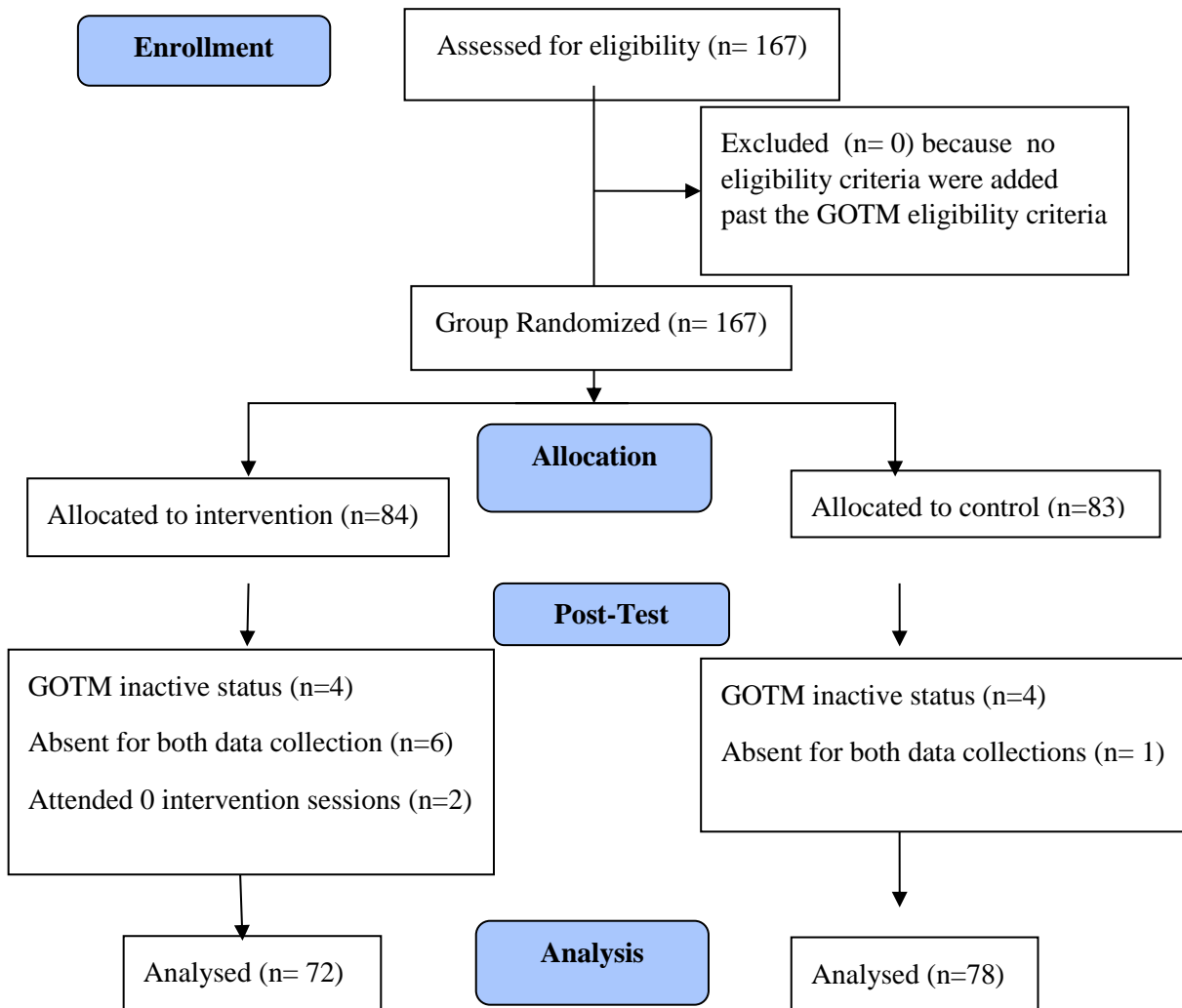


Figure 2 CONSORT Diagram of Loss and Exclusions

The sample consists of females between the ages of 10 and 15. The average age of participants was 12 years and the majority of participants were in 6<sup>th</sup> and 7<sup>th</sup> grade (83.8%). The majority of participants were eligible for free and reduced lunch (81.4%). The majority of guardians identified the participants as Black or African American (56.3%), followed by White or Caucasian and then Multi-Racial. Tables 3 and 4 below present information about the demographic nature of the sample. These sample characteristics were compared to the demographic information of the four participating schools using a chi square goodness of fit test. Table 5 below presents the demographic information for the four schools from which the sample was drawn. No significant differences were found between the racial composition of the sample and the four school populations.

Table 3  
Descriptive Statistics for Categorical Demographic Variables

Demographic Measure		Frequency	Percent
Race	Asian	10	6.9
	Native Hawaiian or Pacific Islander	0	0
	Alaskan Native or Native American	0	0
	Black or African-American	81	56.3
	White or Caucasian	32	22.2
	Other	3	2.1
	Multi-Racial	18	12.5
Ethnicity	Hispanic	26	18.3
	Non-Hispanic	116	81.7
Free and Reduced Lunch	Eligible	114	81.4
	Not Eligible	27	18.6
Grade	5 <sup>th</sup>	16	10.4
	6 <sup>th</sup>	61	39.6
	7 <sup>th</sup>	68	44.2
	8 <sup>th</sup>	9	5.8

Table 4

## Descriptive Statistics for Measures Prior to Maximum Likelihood Estimation of Missing Data

Measure	Mean	SD	Skewness	Kurtosis	% Missing
Age	12.1	1.03	0.25	0.04	0
Days Absent from School	4.8	4.6	2.1	7.0	9.3
Percent Club Attendance*	50.6	29.3	0.1	0.6	0
Pre-Test WJMF	92.1	12.9	.193	-0.10	16.7
Post-Test WJMF	96.3	48.8	10.4	116.0	9.3
Pre-Test WJAP	85.7	18.6	-2.3	8.3	15.3
Post-Test WJAP	91.6	53.8	10.7	122.1	8.7
Pre-Test NIH FIT	104.8	12.0	-0.5	-0.2	10.0
Post-Test NIH FIT	109.5	12.4	-0.8	0.5	9.3
Pre-Test NIH DCCS	103.2	15.2	-2.0	9.3	13.3
Post-Test NIH DCCS	107.3	12.1	-1.2	2.2	11.3
Pre-Test NIH LS	93.7	13.1	-0.1	-0.6	12.7
Post-Test NIH LS	96.9	13.8	0.2	0.4	10.0
Pre-Test BMI	22.9	6.0	1.1	0.9	2.7
Post-Test BMI	23.5	6.4	1.2	1.2	6.7
Pre-Test Body Fat %	29.4	10.1	0.4	-0.6	0.7
Post-Test Body Fat %	29.8	10.0	0.2	0.4	3.3
Pre-Test MVPA Min/Hour	3.0	1.1	1.1	2.1	2.0
Post-Test MVPA Min/Hour	3.0	1.2	0.4	-0.5	7.3
Pre-Test VO2 MAX Max	38.2	5.1	-1	0.7	2.0
Post-Test VO2 MAX Max	37.8	5.5	-1.1	1	6.7

*Note.* \*Intervention group only

The data were analyzed for group differences between the control and intervention groups on descriptive and pre and post-test measures. Some measures were not normally distributed for both or one of the groups, as assessed by Shapiro-Wilk's test ( $p < .05$ ). Due to the non-normal distribution of age and school attendance data, BMI, VO<sub>2</sub> max, MVPA, applied

problems, shifting, and inhibition, the non-parametric Kruskal-Wallis H test was used. The groups were not statistically significantly different on any of these variables, which indicates that at baseline these groups did not differ in their age, rates of attendance at school, body fat percentage, BMI, cardiovascular fitness, math problem solving, math fluency, shifting, working memory, or inhibition ( $p > .05$ ).

Table 5  
District Demographic information from the Performance and Information Single Record Student Database (SRSD)

District	<i>Kentwood</i>		<i>Lansing</i>	
	School 1	School 2	Academy 1	Academy 2
School Building				
Free and Reduced Lunch	69%	52.90%	75.50%	78.90%
Female	48.30%	48.20%	42.70%	45.20%
Asian	17.40%	10%	5.70%	9.95%
African-American	39.10%	24.40%	59.30%	47.50%
White	26.80%	44.10%	15.40%	20.10%
Hispanic	10%	12.20%	13%	16.70%
Multi-Racial	6.80%	8.30%	6.10%	4.70%

Note: Data are for students in grades 5<sup>th</sup> through 8<sup>th</sup> grade. Ethnicity percentages are the percentage of females in those categories out of total female students.

However, intervention and control groups may have meaningful differences in their racial composition and rate of free and reduced lunch status. There was a statistically significant association between intervention group and reported racial group,  $\chi^2(4) = 10.24$   $p = 0.037$ . There was also a statistically significant association between intervention group and free and reduced lunch eligibility,  $\chi^2(1) = 8.308$   $p = 0.004$ . The intervention group had significantly fewer students who identified as White or Caucasian and significantly more students who identified as Black or African-American. 11.6% of the intervention group identified as White or Caucasian and 66.7% identified as Black or African American while 33.3% and 47.2% of the control group identified

as these groups respectively. Observed and expected counts, percentages of the intervention groups, and adjusted residuals are presented below in Table 6 and Table 7. The strength of the association between intervention group status and both variables however was small (Cohen, 1988), Cramer's  $V < .30$ . Therefore, the intervention and control groups differed significantly on these variables but the strength was small.

Table 6  
Crosstabulation of Racial Category and Intervention Group

		Race				
		Asian	Black/ African- American	White/ Caucasian	Other	Multi- Racial
Control	Count	4 <sub>a,b</sub>	34 <sub>b</sub>	24 <sub>b</sub>	2 <sub>a,b</sub>	8 <sub>a,b</sub>
	Expected	4.6 (-.4)	40.9 (-2.3)	16.3 (3.1)	1.5 (.5)	8.7 (-.4)
	%	5.6	47.2	33.3	2.8	11.1
Intervention	Count	5 <sub>a,b</sub>	46 <sub>b</sub>	8 <sub>b</sub>	1 <sub>a,b</sub>	9 <sub>a,b</sub>
	Expected	4.4 (.4)	39.1 (2.3)	15.7 (-3.1)	1.5 (.5)	8.3 (.4)
	%	7.2	66.7	11.6	1.4	13.0

Note: Each subscript letter denotes subset whose column proportions do not differ significantly from each other at the .05 level. Adjusted residuals are in parenthesis next to the expected count values.

Table 7  
Crosstabulation of Free and Reduced Lunch Status and Intervention Group

		Free and Reduced Lunch Status	
		No	Yes
Control	Count	20 <sub>a</sub>	52 <sub>b</sub>
	Expected Count	13.4 (2.9)	58.6 (-2.9)
	% of Control	27.8	72.2
Intervention	Count	6	62
	Expected Count	12.6 <sub>a</sub> (-2.9)	55.4 <sub>b</sub> (2.9)
	% of Intervention	8.8	91.2

Note: Each subscript letter denotes subset whose column proportions do not differ significantly from each other at the .05 level. Adjusted residuals are in parenthesis next to the expected count values

## Missing Data

Analysis of the data completion indicates that there was a significant amount of missing data. Of the 3,205 total values examined, 7% were missing. As shown in Table 4, the pre-test measures of math achievement and executive functioning had the highest levels of missing data. Club attendance and age data have the lowest level of missing data, with all cases being complete cases.

Data missing at random is data that is unrelated to the variable itself but can still be related to other measured variables. Data missing completely at random (MCAR) is a more stringent missing data pattern where the likelihood of missing data is unrelated to other measured variables and to that variable itself. Little's MCAR analysis was conducted to examine the possibility that the data are missing completely at random. Little's MCAR test was not significant ( $\chi^2(613, N = 150) = 579.75, \text{Sig.} = 0.83$ ), suggesting that data in the sample may be missing completely at random. Analysis of the data patterns indicate that there are only two patterns that occur in more than 1% of the cases. The first pattern is of participants missing pre-test inhibition math fluency and applied problems. The second pattern is of participants missing all three pre-test executive function measures. These patterns suggest that the missingness of the variables is related to missing pre-test measurement data collection dates. Missingness of data due to unexpected personal events that cause absences or scheduling difficulties are considered in line with data missing completely at random. Together this information suggests, although is not definitive proof, that the data are missing completely at random. In addition, a violation of the missing at random assumption that there is no relation between a variable and the likelihood of missing data on that variable is considered problematic when an unmeasured cause of missingness has a strong relationship with the variable even after partialing out other measured



variables. This situation is considered unlikely by leading statisticians. Together, this indicates the appropriateness of the use of Maximum Likelihood Estimation (MLE) as a method for dealing with missing data as MLE methods relies on the assumption that data are missing at random. For this study, MLE was completed using SPSS's regression method and estimation adjustment with residuals. This method estimates missing values using multiple linear regression where error terms chosen randomly from the observed residuals of complete cases are added to the regression estimates.

### **Intervention Fidelity**

Participants in the GOTM afterschool program attended an average of 50.6% (26.28) of the 49-50 intervention sessions provided. The median (51%) percentage of sessions attended was similar. The modes for percentage of sessions attended were 16% and 88% of sessions. The number of sessions attended ranged from 1 to 50 sessions, which is a range of 2 to 100% of sessions. Club attendance for each participant was also converted into a categorical variable with four levels. The levels were 0 to 25% attendance, 25 to 50% attendance, 50 to 75% attendance, and 75% or greater attendance. Of the 72 participants in the intervention condition, 18 attended less than 25% sessions, 16 attended 25 to 50%, 20 attended 50 to 75%, and 18 attended 75 to 100%. A chi-square goodness-of-fit test was conducted to determine whether there were an equal number of participants at each attendance level. The minimum expected frequency was 18. The chi-square goodness-of-fit test indicated that the attendance levels were equally represented by the participants ( $\chi^2(3) = 0.444, p = .931$ ). Therefore, there were not significantly less low attending participants than average or high attending participants.

Information on the following fidelity checklist items was analyzed for this study: time spent in seated or standing activities, time spent on management of the club, time spent on

instruction, total club time, and step counts and MVPA from girls wearing accelerometers during the session. Ten fidelity observations were completed at both intervention schools. Descriptive statistics of these measures are provided below in Table 8. This information was compared to the Intervention Protocol plan. The mean of 88.19 total club minutes was close to the goal of 90 minutes in the protocol. Both the mean minutes of observed opportunities for MVPA (16.88 minutes) and the mean minutes spent in MVPA (21.52 minutes) were lower than the intervention protocol plan of 60 minutes of opportunity for MVPA.

Table 8  
Descriptive Statistics of Intervention Fidelity Checklist Measures

	N	Minimum	Maximum	Mean	SD
Total Club Time (min)	20	80	94	88.19	4.002
Time in Seated/Standing Activities (min)	20	0	19	7.30	5.83
Time spent on Management (min)	20	11	37	23.74	6.79
Time spent on Instruction (min)	20	1	10	5.04	2.62
Observed Opportunity for MVPA	20	5	32	16.88	8.33
MVPA (avg. min)	20	5	33	21.52	6.80
Step Counts (avg.)	20	917	4363	2863.97	897.71

Note: N= number of observations completed

To examine potential school level differences in intervention fidelity, both attendance and intervention fidelity checklist measures were utilized. Looking at the school level, in Kentwood the number of sessions attended ranged from 8% to 100% of sessions with a mean of 57.5% of sessions. In Lansing, the number of sessions attended ranged from 2 to 95%, with a mean of 43.3% of sessions. An independent-samples t-test was run to determine if these were significant differences. There was homogeneity of variances, as assessed by Levene's test for equality of variances ( $p = .697$ ). Level of attendance at the club, as measured by percentage of club sessions attended, was higher for Kentwood ( $M = 57.5\%$ ,  $SD = .29$ ) than for Lansing ( $M = 43.3\%$ ,  $SD = .282$ ), a statistically significant difference,  $M = .141$ , 95% CI [0.006, 0.276],  $t(70) = 2.096$ ,  $p =$

.04. Therefore, participants at Crestwood attended 14% more sessions on average than participants at Gardner.

Next, chi-square test of independence was conducted comparing the frequencies of intervention attendance levels in the different schools (0, 25, 50, and 75%). The number of students who attended 75% or more of sessions appeared higher at Crestwood (32.4% of participants) than at Gardner (8.3%). However, there was no statistically significant association between school and level of attendance,  $\chi^2(3) = 3.369, p=0.338$ . Therefore, the schools did not significantly differ in the number of students who fell into each level of attendance, see Table 9. Together with the descriptive data and independent samples t-test, this suggests that while the schools differ in the average percentage of sessions attended, the schools had similar levels of students in each attendance level.

Table 9  
Results of Chi-square Test and Descriptive Statistics for Club Attendance Level by School

Level of Club Attendance	District	
	Kentwood	Lansing
0 to 25% sessions attended	8 (21.6)	10 (25)
25 to 50% sessions attended	6 (16.2)	10 (28.6)
50 to 75% sessions attended	11 (29.7)	9 (25.7)
75% or more sessions attended	12 (32.4)	6 (8.3)

Count followed by percentage of the sample at that school in parenthesis. \* $p < .05$ .  $df = 2$ .

Analysis was completed to determine if any school level differences existed in the fidelity checklist measures by intervention school. Data were normally distributed for each group, as assessed by Shapiro-Wilk test ( $p > .05$ ); and there was homogeneity of variances for time spent in seated/standing activities, minutes of management, minutes of instruction, and observed opportunities for MVPA as assessed by Levene's test of homogeneity of variances ( $p > .001$ ). Levene's test indicated unequal variances for actual MVPA, step counts and total club time, therefore equal variance was not assumed for those measures and a one-way Welch

ANOVA statistic was utilized for those variables. Based on a one-way ANOVA analysis, the differences between the two schools on time spent in seated/standing activities,  $F(1, 18) = 0.239$ ,  $p = .631$ , minutes of management,  $F(1, 18) = 0.738$ ,  $p = .402$ , minutes of instruction,  $F(1, 18) = 0.035$ ,  $p = .854$ , and observed opportunities for MVPA,  $F(1, 18) = 0.512$ ,  $p = .483$ , were not statistically significant.

The difference between the groups for actual MVPA, Welch's  $F(1, 10.409) = 0.00$ ,  $p = .991$ , and step counts were also not significant, Welch's  $F(1, 11.473) = 0.064$ ,  $p = .805$ . Tests of equality for total club time could not be completed because total club minutes for Gardner had no variance. See Table 10 below for more details. Therefore, the two school clubs did not differ in

Table 10  
Univariate Effects for School on Club Fidelity Measures

	F/Welch	p	School	Means	99.9% CI
Time Seated/Standing (min)	0.239	0.63	Kentwood	6.65	2.7 – 10.60
			Lansing	7.95	4.0-11.9
Time spent on Management (min)	0.738	0.40	Kentwood	22.43	17.89-26.96
			Lansing	25.05	20.51-29.59
Time spent on Instruction (min)	0.035	0.85	Kentwood	5.2	3.36-6.94
			Lansing	4.93	3.14-6.71
Observed Opportunity for MVPA	0.512	0.48	Kentwood	18.23	12.62-23.83
			Lansing	15.53	9.92-21.13
MVPA (avg. min)	0.000	0.99	Kentwood	21.51	18.87-26.14
			Lansing	21.54	16.9-26.18
Step Counts (avg.)	0.064	0.81	Kentwood	2811.86	2200.2-3423.5
			Lansing	2916.08	2304.4-3517.7

Note: df, error df. is (1, 18). 99.9% CI = confidence interval reporting lower- upper bound. these measures of intervention fidelity.

## Changes in Cardiovascular Fitness, Body Composition, and Physical Activity.

Group means with 95% confidence intervals at pre-test and post-test and the change values are provided in Table 11. The analysis presented below includes the entire sample, as the models run using samples selected based on intervention participants attending at least 25% of GOTM sessions or at least 50% of GOTM sessions did not differ in terms of significance ( $p > .05$ ).

Table 11  
Means and 95% Confidence Interval for groups at Pre and Post-test and Change Values

	Intervention			Control		
	Pre-Test	Post-Test	Change	Pre-Test	Post-Test	Change
VO <sub>2</sub> Max	38.1 (36.8-39.3)	37.6 (36.4-38.9)	-0.5	38.1 (36.9-39.2)	37.8 (36.6- 39.0)	-0.9
Body Fat Percentage	29.4 (27.0-31.7)	29.9 (27.6-32.2)	0.5	29.5 (27.2-31.7)	29.7 (27.5-32.0)	0.2
BMI (z score)	0.897 (0.6-1.2)	0.931 (0.7-1.2)	-0.06	0.884 (0.6-1.1)	0.925 (0.7-1.2)	0.04
MVPA	2.96 (2.7-3.2)	2.97 (2.7-3.3)	0.01	3.05 (2.8-3.3)	3.15 (2.9-3.4)	0.1

Analyses of the effect of the GOTM intervention on cardiovascular fitness, BMI, Body fat percentage, and MVPA was conducted by using 2 (group: intervention, wait-list) by 2 (time: pretest, posttest) repeated measures analysis of variance (ANOVA). A multivariate analysis was not appropriate given multicollinearity between many of the variables ( $r > .9$ ). A complete table of Correlations between the variables is presented in Appendix C.

VO<sub>2</sub> max data, the measure of cardiovascular fitness, was normally distributed, as assessed by Q-Q plot. There was homogeneity of variances for both pretest ( $p = .815$ ) and post-test VO<sub>2</sub> max values ( $p = .792$ ), as assessed by Levene's test of homogeneity of error variances. The assumption of homogeneity of covariance matrices was not met, as assessed by Box's M test

( $p < .001$ ). Attempts to remove the violation assumption via data transformation were unsuccessful. Therefore, the decision was made to run the analysis with the violation and to follow up with an independent samples t-test on change scores. There was no statistically significant interaction effect of time and intervention group on body fat percentage,  $F(1, 148) = 0.701, p = .40$ . There was however a statistically significant effect of time,  $F(1, 148) = 8.773, p = .004, \eta^2 = .056$ . There was heterogeneity of variances, as assessed by Levene's test for equality of variances ( $p = .004$ ) therefore the analysis below is with equal variances not assumed. The intervention group change score did not differ significantly from the change scores of the control group,  $t(148) = 0.837, p = .404$ . Therefore, the effect of the GOTM intervention on improvement in cardiovascular fitness was not significant in this subsample of the GOTM project.

Table 12  
Repeated measures ANOVAs for VO2 Max, MVPA, BMI, and Body Fat Percentage

		MS	Df	F	p	Partial $\eta^2$
VO2 Max	Intervention	0.647	1, 148	0.012	0.91	0.000
	Time	8.049	1, 148	8.773	0.004*	0.056
	Time * Intervention	0.643	1, 148	0.701	0.40	0.005
Body Fat Percentage	Intervention	0.031	1, 148	0.000	0.99	0.000
	Time	12.654	1, 148	1.304	0.26	0.009
	Time * Intervention	1.332	1, 148	0.137	0.71	0.001
BMI	Intervention	0.007	1, 148	0.003	0.96	0.000
	Time	0.109	1, 148	2.837	0.09	0.019
	Time * Intervention	0.001	1, 148	0.024	0.88	0.000
MVPA	Intervention	1.675	1, 148	0.762	0.38	0.005
	Time	2.891	1, 148	4.282	0.40	0.002
	Time * Intervention	0.124	1, 148	0.184	0.67	0.002

MS: mean square, Df: degree of freedom, error degrees of freedom. \* indicates significance at  $\alpha = .004$   $\eta^2$ : Partial eta squared

Pre and Post-test body fat percentage was normally distributed, as assessed by Q-Q plot. There was homogeneity of variances for both pretest ( $p=.408$ ) and post-test body fat percentage

max values ( $p=.772$ ), as assessed by Levene's test of homogeneity of error variances. The assumption of homogeneity of covariance matrices was not met, as assessed by Box's M test ( $p < .001$ ). Attempts to remove the violation assumption via data transformation were unsuccessful. Therefore, the decision was made to run the analysis with the violation and to follow up with repeated measures ANOVA analysis in each intervention group. There was no statistically significant interaction effect of time and intervention group on body fat percentage,  $F(1, 148) = 0.137$ ,  $p = .712$ . However, a main assumption of the analysis was violated. The follow up repeated measures analysis completed separately for each group, with no assumption violations, indicated there was no significant effect of time on the control group,  $F(1,77) = 0.559$ ,  $p=.457$  or the intervention group,  $F(1,77) = 0.742$ ,  $p=.392$ . This analysis suggests that the effect of the intervention on body fat percentage was not significant in this subsample of the GOTM project.

Pre and Post-test z-score transformed BMI scores were normally distributed, as assessed by Q-Q plot. There was homogeneity of variances for both pretest ( $p=.583$ ) and post-test BMI values ( $p=.92$ ), as assessed by Levene's test of homogeneity of error variances. The assumption of homogeneity of covariance matrices was met, as assessed by Box's M test ( $p = .828$ ). There was no statistically significant interaction effect of time and intervention group on BMI,  $F(1, 148) = 0.024$ ,  $p = .878$ . There was also no significant main effect of time,  $F(1, 148) = 2.837$ ,  $p = .094$ . This indicates that there was no significant change in BMI from pre-test to post-test, either overall or differentially for different groups, in this subsample of the GOTM project.

Pre and Post-test MVPA data were normally distributed, as assessed by Q-Q plot. There was homogeneity of variances for both pretest ( $p=.450$ ) and post-test BMI values ( $p=.704$ ), as assessed by Levene's test of homogeneity of error variances. The assumption of homogeneity of covariance matrices was met, as assessed by Box's M test ( $p = .459$ ). There was no statistically

significant interaction effect of time and intervention group on MVPA,  $F(1, 148) = 0.024$ ,  $p = .878$ . This indicates that there was no significant change in MVPA from pre-test to post-test, either overall or differently by group, in this subsample of the GOTM project.

Overall, this analysis indicates that there are no differential effects of the intervention on cardiovascular fitness (VO2 Max), MVPA, body fat percentage, or BMI in this subsample of the GOTM study.

### **Research Question 1**

Question: Do students in the intervention group have higher executive functioning skills than the control group as the result of an after-school physical activity intervention?

*Hypothesis 1:* Students in the intervention group will have significantly better performance on the tasks of inhibition and working memory than the control group. Performance on the measure of shifting will not be different between the two groups.

Means and change scores for the NIH measures are provided in Table 13. Planned analysis for this research questions was a MANCOVA that included pre-test measures of the three executive functioning measures (shifting, inhibition, and working memory) as well as measures of pre-test body composition and MVPA. These variables did not meet the assumptions of linearity or significant regression relationships required for the (M)ANCOVA and therefore were not entered into the model. While pre-test shifting, inhibition, and working memory and pre-test body fat percentage did meet the linearity and significant regression assumptions, they



violated the homogeneity of regressions assumption, as determined by a significant interaction term of the measures with the independent variable of group. There was homoscedasticity and homogeneity of variances, as assessed by visual inspection of a scatterplot and Levene's test of homogeneity of variance for shifting ( $p=.151$ ), inhibition ( $p=.912$ ), and working memory ( $p=.540$ ) respectively. There was also homogeneity of covariances, as assessed by the multivariate Box's M Test. The observed covariance matrices of the dependent variables are equal across groups ( $p=.282$ ). The dependent variables were all significantly correlated but the assumption of absence of multicollinearity between the dependent variables was met based on Tabachnick & Fidell's (2012) suggestion that no correlation between the dependent variables should be above  $r = .90$ . However, standardized residuals for the overall model were not normally distributed for shifting and inhibition, as assessed by the Shapiro-Wilks test ( $p < .001$ ). Standardized residuals were not normally distributed for shifting intervention group ( $p<.001$ ) and inhibition for either the intervention ( $p=.042$ ), or control groups ( $p=.004$ ) as assessed by the Shapiro-Wilks test. Transformations of the data were not able to resolve the violations of the normality and homogeneity of regressions assumptions.

Table 13  
Means and 95% Confidence Interval of Executive Function Variables for Groups at Pre and Post-test and Change Values

	Intervention			Control		
	Pre-Test	Post-Test	Change	Pre-Test	Post-Test	Change
LS	93.8 (90.5-97.1)	97.4 (94.2-100.6)	3.6	94.8 (92-97.7)	96.7 (93.5-99.8)	1.9
FIT	103.4 (100.4-106.3)	109.6 (106.7-112.5)	6.2	106 (103.4-108.6)	108.5 (105.6-111.4)	2.5
DCCS	101.2 (97.7-104.6)	106.8 (106.7-112.5)	5.6	105.5 (102.8- 108.1)	107.9 (105.5-110.2)	2.4

Because of the violation of multiple assumptions for MANCOVA, analysis of the effect of intervention on the NIH executive function measures was instead conducted using repeated measures analysis of variance (MANOVA). The correlations between working memory and shifting at pre-test ( $r = .209$ ) and post-test ( $r = .178$ ) and between post-test working memory and inhibition ( $r = .243$ ) were weak. Therefore, they were not entered into a single multivariate analysis. In addition, neither MVPA nor body fat percentage were included in the models presented below. The variables did not contribute significantly when included and therefore are not included here in order to present a parsimonious model. Analysis of the effect of intervention on working memory, inhibition, and shifting were conducted by using 2(intervention, control) by 2 (pretest, posttest) or repeated measures analysis of variance (ANOVA) tests. The data were normally distributed, as assessed by Q-Q plot. There was homogeneity of variances for both pretest and post-test working memory values ( $p < .05$ ), as assessed by Levene's test of homogeneity of error variances. The assumption of homogeneity of covariance matrices was met, as assessed by Box's M test ( $p = < .05$ ).

There was no statistically significant interaction effect of time and intervention group on working memory,  $F(1, 148) = 0.598$ ,  $p = .441$ . There was a statistically significant main effect of time on working memory,  $F(1, 148) = 5.46$ ,  $p = .021$ . There was no significant interaction of time and intervention for working memory,  $F(1, 148) = 0.60$ ,  $p = .931$ . Therefore, working memory improved significantly over time, with a small effect size, but that improvement over time did not differ based on participation in the GOTM intervention. There was a statistically significant interaction effect of time and intervention group on inhibition,  $F(1, 148) = 5.054$ ,  $p = .026$ ,  $\eta^2 = .036$ . There was also a statistically significant main effect of time on inhibition,  $F(1, 148) = 27.329$ ,  $p < .001$ ,  $\eta^2 = .156$ . Therefore, inhibition performance improved significantly

over time but that change was different depending on participation in the GOTM intervention. Effect size measures indicate that 15.6% of the variance in inhibition scores is attributed to the effect of the time and intervention interaction. There was no statistically significant interaction effect of time and intervention group on shifting,  $F(1, 148) = 3.221, p = .075$ . There was a statistically significant main effect of time on working memory,  $F(1,148) = 19.039, p < .001, \eta^2 = .114$ . Therefore, shifting performance improved significantly over time, but that improvement over time did not differ based on participation in the GOTM intervention. Details of the repeated measures ANOVAs can be found below in Table 14.

Table 14  
Repeated Measures ANOVAs for Working Memory, Inhibition, and Shifting

		MS	Df	F	p	Partial $\eta^2$
Working Memory	Time	545.06	1, 148	5.46	0.021*	0.036
	Intervention	1.69	1, 148	0.01	0.931	0.000
	Time * Intervention	59.66	1, 148	0.60	0.441	0.004
Inhibition	Time	1427.24	1, 148	27.33	<0.001**	0.156
	Intervention	45.54	1, 148	0.18	0.672	0.001
	Time * Intervention	263.96	1, 148	5.05	0.026*	0.033
Shifting	Time	1191.24	1, 148	19.04	<0.001**	0.114
	Intervention	545.94	1, 148	2.058	0.154	0.014
	Time * Intervention	201.54	1, 148	3.22	0.075	0.021

MS: mean square, Df: degree of freedom, error degrees of freedom. \*  $p < .05$  \*\*  $p < .01, \eta^2$ : Partial eta squared

To follow up on the significant effect on inhibition, differences in improvement by group were assessed by an independent samples t-test on change scores. There was homogeneity of variances, as assessed by Levene's test for equality of variances ( $p = .874$ ). The intervention group improved significantly more ( $6.2 \pm 9.7$ ) than the control group ( $2.5 \pm 10.6$ ), a statistically significant difference of 3.8 more standard score points (mean = 100, standard deviation = 15) improvement,  $t(148) = -2.248, p = .026, \text{Cohens } d = .37$ . Therefore, the effect of the GOTM

intervention on improvement in inhibition is statistically significant but small. According to Cohen (1998), this is a small effect size. Analysis using Cohen's U3 (1998) determines an improvement index of 15.5, which is interpreted to mean that a control student at the 50<sup>th</sup> percentile could be expected to increase to the 65<sup>th</sup> percentile by participating in the intervention group.

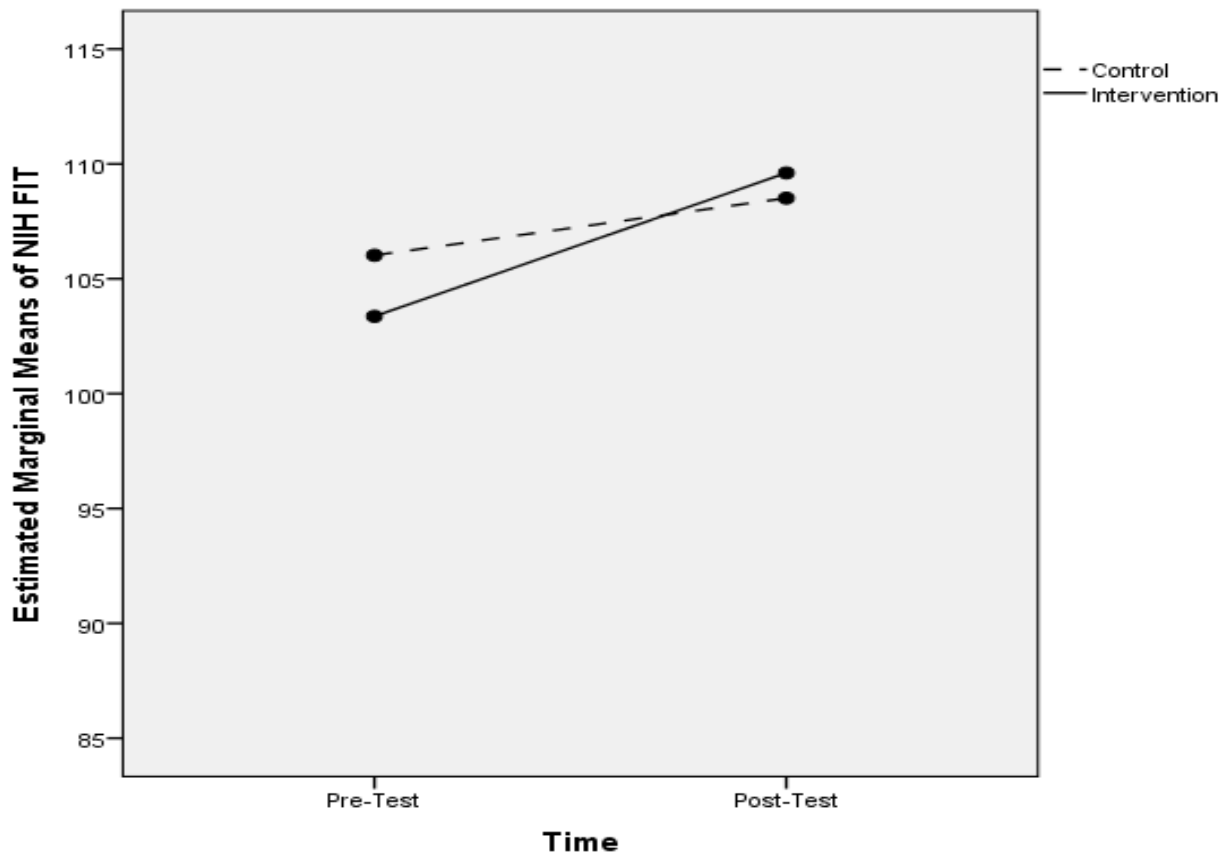


Figure 3 Profile Plot of Intervention and Control Group Marginal Means on NIH FIT at Pre-test and Post-test.

Overall, the hypotheses that students in the intervention group would have significantly more improved performance on the task of inhibition and that performance on the measure of shifting will not be different between the two groups were supported. The hypothesis that the intervention group would have more improved performance on the measure of working memory

was not supported (NIH LS). The effect size of the physical activity intervention on inhibition was found to be small in size.

## **Research Question 2**

Question: Do students in the intervention group have higher math performance than the control group as the result of an after-school physical activity intervention?

*Hypothesis 2:* Significant differences between groups will be seen in both academic measures assessed (math fluency and applied problems), with the intervention group showing more improvements than the control group.

To examine the second research question, one-way MANCOVAs were conducted to assess if differences exist on the math achievement measures of post-test applied problems and post-test math fluency by the independent variable of intervention group (intervention vs. control). Table 15 below reports the overall results for the MANCOVA. Both pre-test math achievement (applied problems and math fluency) were entered as covariates because they met the linearity assumption of the MANCOVA as assessed by visual inspection of a scatterplot, had significant regression values with the dependent variables, and met the assumption of homogeneity of regressions as assessed by examining the significance of the interaction term. While pre-test measures of body composition and MVPA had been proposed as potential covariates, none of these variables met the assumptions of linearity or significant regression relationships required for the MANCOVA. In addition, neither measure had a significant effect on the MANCOVA model when entered ( $p < .05$ ) and the variance explained did not change ( $R^2 =$

.76 in both models for math fluency and  $R^2=.63$  in both models for applied problems). Therefore, to present a parsimonious model they are not included in the following analysis.

Standardized residuals for the intervention groups and for the overall model were normally distributed, as assessed by Kolmogorov-Smirnov test ( $p > .05$ ). There was homoscedasticity and homogeneity of variances, as assessed by visual inspection of a scatterplot and Levene's test of homogeneity of variance for applied problems ( $p = .101$ ) and math fluency ( $p=.120$ ) respectively. There was also homogeneity of covariances, as assessed by the multivariate Box's M Test. The observed covariance matrices of the dependent variables are equal across groups ( $p=.228$ ). The assumption of absence of multicollinearity between the dependent variables was met based on Tabachnick & Fidell's (2012) suggestion that no correlation between the dependent variables should be above  $r = .90$ . Please see Appendix C for all variable correlations.

Table 15  
MANCOVA table for Post-Test Math Fluency (WJMF) and Problem Solving (WJAP) by Intervention and Controlling for Pre-test WJMF and WJAP

Dependent Variable	Variable	df	F	p	$\eta^2$
Post-Test WJMF	Pre-test WJMF	1,146	319.32	<.001**	.686
	Pre-test WJAP	1,146	7.75	.005**	.05
	Intervention	1,146	8.41	.004**	.054
Post-Test WJAP	Pre-test WJMF	1,146	3.89	.05*	.036
	Pre-test WJAP	1,146	171.41	<.001**	.054
	Intervention	1,146	.382	.538	.003

MS: mean square, SD: Standard Deviation, Df: degree of freedom, error degrees of freedom. \*  $p<.05$  \*\*  $p<.01$

After controlling for the covariates of pre-test math fluency and applied problems, there was a statistically significant difference in the overall model based on intervention condition,  $F(2, 145) = 4.183, p = .017$ ; Wilk's  $\Lambda = 0.945$ , partial  $\eta^2 = .055$ . Because the multivariate analysis

is significant, post hoc analysis consisting of ANCOVA with the two dependent variables separately was conducted to assess specific differences among the variables by intervention group. Both pre-test math fluency and math problem solving were significantly associated with post-test math fluency. After adjustment for pre-test math fluency and applied problems scores, there was a statistically significant difference in post-intervention math fluency scores between the intervention groups,  $F(1, 146) = 8.406$ ,  $p = .004$ , partial  $\eta^2 = .054$ .

In line with the analysis required for the first research questions, analysis of the effect of intervention on math fluency and applied problems was also conducted by using a 2 (intervention, control) by 2 (pretest, posttest) multivariate analysis of variance (MANOVA). The data were normally distributed, as assessed by Q-Q plot. There was no multicollinearity, as all Pearson correlation values are below 0.9. There was homogeneity of covariances ( $p = .433$ ), as assessed by Box's M test. There was a statistically significant interaction effect between time and intervention group on the combined dependent variables,  $F(2, 147) = 4.561$ ,  $p = .01$ ,  $\eta^2 = .058$ . Univariate tests indicated there was a time by intervention effect on math fluency,  $F(1, 148) = 8.845$ ,  $p = .003$ ,  $\eta^2 = .056$ . According to the partial eta squared statistic 5.6% of the total variance explained by the error and the time by intervention effect is due to the time by intervention effect. This is a small effect size. The time by intervention effect was not significant for applied problems,  $F(1, 148) = 0.584$ ,  $p = .446$ . There were no significant main effects of time or intervention. This indicates that math fluency performance from pre to post-test changed differently depending on participation in the GOTM intervention. Figure 6 below provides a visual representation of the differences in scores over time by group.

Table 16

Repeated Measures ANOVAs for Math Fluency and Math Problem Solving

		MS	Df	F	p	Partial $\eta^2$
WJMF	Time	0.06	1,148	.003	.958	0.000
	Intervention	104.32	1,148	.315	.575	0.002
	Time * Intervention	203.81	1,148	8.845	.003**	0.056
WJAP	Time	35.72	1,148	1.014	.316	0.007
	Intervention	144.62	1,148	.471	.494	0.003
	Time * Intervention	20.59	1,148	.584	.446	0.004

MS: mean square, Df: degree of freedom, error degrees of freedom, \* p<.05 \*\* p<.01,  $\eta^2$ : Partial eta squared

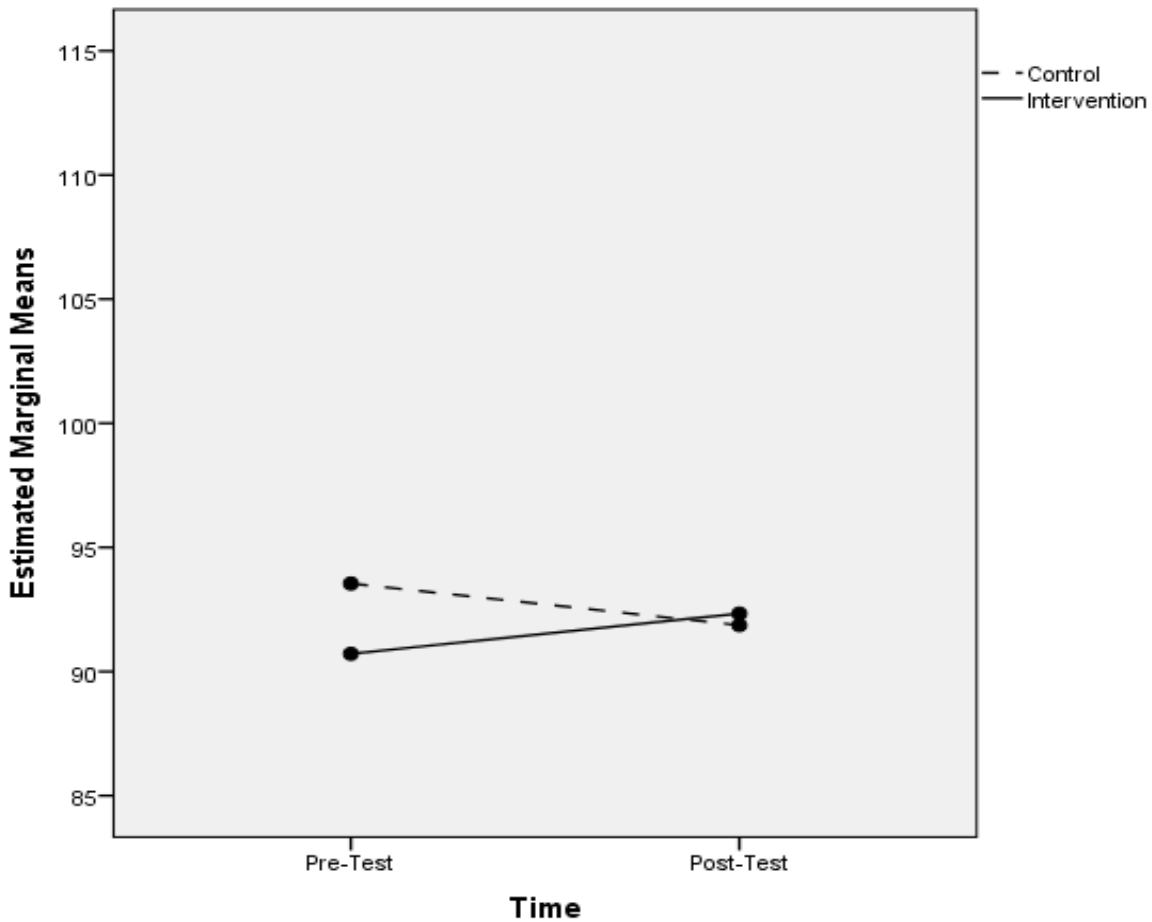


Figure 4. Profile Plot of Intervention and Control Group Marginal Means on Math Fluency (WJMF) at Pre-test and Post-test



Table 17

Means and 95% Confidence Interval of Math Achievement Measures for Groups at Pre and Post-test and Change Values

	Intervention			Control		
	Pre-Test	Post-Test	Change	Pre-Test	Post-Test	Change
WJMF	90.7 (87.4-94)	92.3 (88.8-95.9)	1.6	93.5 (90.9-96.2)	91.9 (89.1-94.6)	-1.6
WJAP	86.7 (83.5-89.9)	86.5 (83.4-89.6)	-0.2	88.6 (85.6-91.6)	87.4 (84.7-90.1)	-1.2

To follow up, differences in math fluency improvement by group were assessed by an independent samples t-test on change scores. There was homogeneity of variances, as assessed by Levene's test for equality of variances ( $p = .184$ ) so equal variances were assumed in the following analysis. The intervention group improved significantly more ( $1.62 \pm 6.2$ ) than the control group ( $-1.68 \pm 7.3$ ), a statistically significant difference of 3.3 standard score points (95% CI,  $-5.49$  to  $-1.01$ ),  $t(148) = -2.974$ ,  $p = .003$ , Cohens  $d = .49$ . Table 17 presents all results. Therefore, the effect of the GOTM intervention on improvement in math fluency is statistically significant and approaches a medium effect size (Cohen, 1998). Analysis using Cohens U3 (1998) found an improvement index of 19, which is interpreted to mean that a participant in the control group that is at the 50<sup>th</sup> percentile could be expected to increase to the 69<sup>th</sup> percentile if they participated in the intervention. While the control group showed a decrease in performance, the intervention group showed a small increase in performance.

Together, these analyses indicate that the intervention group's post-test math fluency performance is significantly higher than the control group, by close to 3 standard score points. Importantly, the analysis indicates that part of the difference in math fluency performance may come from a decrease in performance in the control group at post-test. Both analyses (MANCOVA and Repeated Measures ANOVA) failed to find a difference in applied problems performance by intervention group. Therefore, the overall hypothesis that significant differences

between groups would be seen in both academic measures assessed (math fluency and applied problems), with the intervention group showing more improvements in math fluency than the control group, was not supported by this analysis.

### **Research Question 3**

Question: What is the relation between change in executive functions and change in math achievement? If there are differential improvements in academic achievement by intervention group, are these differences related to improvements in executive functioning?

*Hypothesis 3a.* Improvements in math problem solving will be related to improvements in working memory and shifting.

*Hypothesis 3b.* Improvements in math fluency will be related to shifting and inhibition.

Preliminary analysis of the relationship between executive function and math achievement was completed using correlation analysis. As shown in Table 18, none of the correlations between the executive function change scores and the math achievement change scores were significant.

Analysis of the mediation effect of executive functions on the impact of intervention on math fluency was conducted using the PROCESS macro for SPSS from Hayes (2016). Specifically, the model 4 specification of multiple simultaneous mediators from the PROCESS macro was used. One model was completed for math fluency and one for math problem solving.

Table 18  
 Pearson Correlations Between Executive Function and Math Achievement Change Scores

	Inhibition Change	Working Memory Change	Shifting Change
Math Fluency Change	0.068	0.080	0.066
Math Problem Solving Change	-0.009	-0.008	-0.018

\* Correlation is significant at the  $p < 0.05$

A multiple simultaneous mediator model was used to estimate changes in inhibition, shifting, working memory and math problem solving from intervention as well to estimate change in math problem solving from intervention status, change in inhibition, change in shifting, and change in working memory. There was a significant effect of intervention on inhibition, as intervention was positively related to change in inhibition ( $a = 3.76, p = .033$ ). Neither change in shifting nor change in working memory was significantly estimated by intervention status which indicates no effect of intervention on these executive functions. None of the change in executive function variables significantly predicted change in math problem solving ( $p > .05$ ). Therefore, there was no direct effect of intervention on math problem solving. The indirect effect of intervention on math problem solving through change in the three executive functions was examined using Hayes (2013) method of bootstrapping confidence intervals and used 10,000 bootstrap samples. There was no evidence for a total indirect effect of the summed executive function variables, CI [-0.916 to 0.684]. There was also no evidence for an indirect effect of change in inhibition, CI [-0.738-0.569], change in shifting, CI [-0.551 – 0.427]. or change in working memory, CI [-0.523 – 0.16]. Table 19 and Figure 5 provide more information.

Table 19

Model Summary, Regression Coefficients, Standard Errors, and Significance Values for the Simultaneous Multiple Mediator Model of Math Problem Solving

	<u>Δ Inhibition</u>			<u>Δ Shifting</u>			<u>Δ WM</u>			<u>Δ Prob. Solving</u>		
	Coeff.	SE	P	Coeff.	SE	P	Coeff.	SE	P	Coeff.	SE	P
Intervention	3.76	1.67	<b>.03</b>	3.28	1.83	.07	1.79	2.31	.44	1.16	1.42	.42
Δ Inhibition	-	-	-	-	-	-	-	-	-	-0.02	0.07	.82
Δ Shifting	-	-	-	-	-	-	-	-	-	-0.01	.06	.88
Δ WM	-	-	-	-	-	-	-	-	-	-0.01	.050	.80
Constant	2.49	1.16	.03	2.35	1.27	.07	1.81	1.6	.26	-1.13	.99	.25
	R <sup>2</sup> = 0.033			R <sup>2</sup> = 0.021			R <sup>2</sup> = 0.004			R <sup>2</sup> = 0.005		
	F(1, 148)= 5.05			F(1, 148)= 3.22			F(1, 148)= 0.60			F(1, 145)= 0.18		
	p = <b>.03</b>			p = .08			p = .44			p = .95		

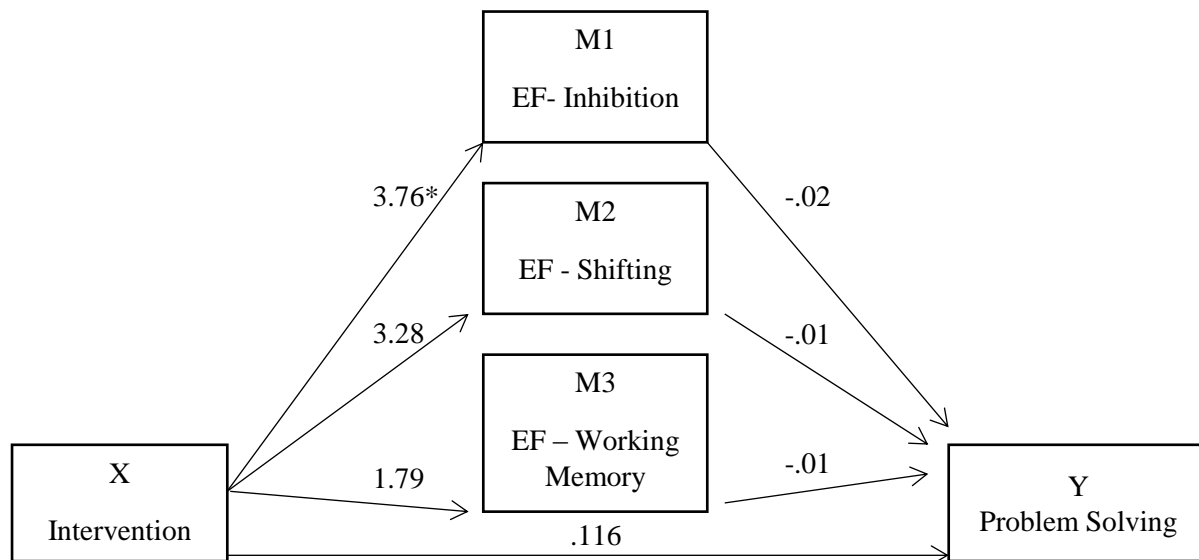


Figure 5 Regression Coefficients for Simultaneous Multiple Mediation Model of Problem Solving. \*indicates significance at  $p < .05$

A multiple serial mediator model was also used to estimate change in math fluency from interventions status, change in inhibition, change in shifting, and change in working memory. There was a significant direct effect of intervention status on math fluency ( $a = 3.13, p = .01$ ). The indirect effect of the intervention on math problem solving through change in the three executive functions was examined using Hayes (2013) method of bootstrapping confidence

interval based on 10,000 bootstrap samples. There was no evidence for a total indirect effect of the summed executive function variables, CI [-0.488 to 1.006]. There was also no evidence for an indirect effect of change in inhibition, CI [-0.469-0.588], change in shifting, CI [-0.168 – 0.625]. or change in working memory, CI [-0.105 – 0.548]. Table 20 and Figure 6 below provide more information.

Table 20  
Model Summary, Regression Coefficients, Standard Errors, and Significance Values for the Simultaneous Multiple Mediator Model of Math Fluency

	<u>Δ Inhibition</u>			<u>Δ Shifting</u>			<u>Δ WM</u>			<u>Δ Fluency</u>		
	Coeff.	SE	P	Coeff.	SE	P	Coeff.	SE	P	Coeff.	SE	P
Intervention	3.76	1.67	<b>.03</b>	3.28	1.83	.07	1.79	2.31	.44	3.13	1.15	<b>.01</b>
Δ Inhibition	-	-	-	-	-	-	-	-	-	.014	.056	.80
Δ Shifting	-	-	-	-	-	-	-	-	-	.024	.051	.64
Δ WM	-	-	-	-	-	-	-	-	-	.023	.04	.55
Constant	2.49	1.16	<b>.03</b>	2.35	1.27	.07	1.81	1.6	.26	-1.81	.80	<b>.02</b>
	R <sup>2</sup> = 0.033			R <sup>2</sup> = 0.033			R <sup>2</sup> = 0.033			R <sup>2</sup> = 0.061		
	F(1, 148)= 5.05			F(1, 148)= 5.05			F(1, 148)= 5.05			F(1, 145)= 2.36		
	p = <b>.03</b>			p = .03			p = .03			p = <b>.05</b>		

This information provides additional partial support for the part of hypothesis which predicted that the intervention group would show more change in inhibition. In line with previous analysis of research question 2, Hypothesis 2 is again partially supported by these results which also indicate an intervention effect on math fluency but not math problem solving. Hypothesis 3a that the change in math problem solving seen as a result of the intervention will be mediated by the change in any of these executive functions is not supported by these results. Hypothesis 3b that the change in math fluency seen as a result of the intervention will be mediated by the change in any of these executive functions is also not supported. In summary, participation in the intervention was positively associated with level of change in inhibition and

math fluency however no mediation effect of any executive functions, together or separately, was found for either math fluency or math problem solving.

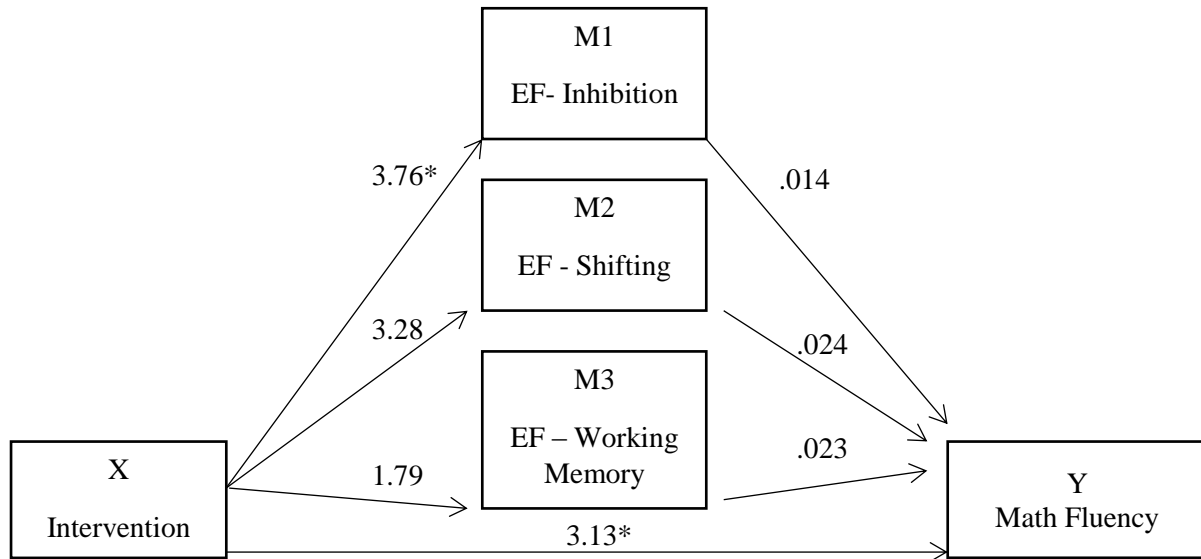


Figure 6 Regression Coefficients for Simultaneous Multiple Mediation Model of Math Fluency. \*indicates significance at  $p < .05$

## CHAPTER 5: DISCUSSION

Regular participation in physical activity is an acknowledged part of a healthy lifestyle and can have significant health effects, such as reduced risk for many diseases and overall improvement in the health of muscles, bones, and joints (Department of Health and Human Services and Department of Education, 2000; U.S. Department of Health and Human Services, 2008). However, research has indicated that current generations of children are less active than previous generations (Boreham & Riddoch, 2001; Burkhalter & Hillman, 2011). There is evidence that level of physical activity is an important factor in brain health and cognition, particularly executive function, in older adult populations. If physical activity is also a key factor in brain health and cognition in children, this would be important to understand as physical activity may then serve as an important factor for intervention during the important childhood years where extensive changes in brain structure and connectivity occur. In addition, because executive function has been shown to be related to certain aspects of academic achievement, it is possible that physical activity may be related to academic achievement. If so, physical activity may serve as a point of intervention for supporting cognitive development and academic achievement as well as its current role as an intervention for physical health.

Despite this importance, this area is currently understudied relative to research in older adult populations. In particular, the specific effects of physical activity on key cognitive process such as executive function and its effects on the important childhood outcome of academic performance remain unclear. Therefore, the purpose of this study was to test the hypothesis that an after school physical activity intervention will increase academic achievement in math and executive functioning in adolescent girls living in low-income communities and to examine how the level of change in executive functions related to change in math achievement. These

questions were examined by randomly assigning female participants in grades 5 to 8 at the school level to an after-school club providing physical activity opportunities or to offerings as usual and assessing their math performance and the executive functions of inhibition, shifting, and working memory at both pre-test and post-test. Intervention participants improved significantly more in inhibition and math fluency, however no mediation effect of inhibition or other executive function change on change in math fluency or math problem solving was seen.

### **Results of Intervention on Physical Activity Measures**

Results indicate that intervention fidelity and dosage were low and that changes in measures of physical activity, cardiovascular fitness and body composition were not seen. Attendance at the club, on average, was low and the number of participants attending fewer than a quarter of sessions was not different from the number attending half or more. The club sessions were held for the planned amount of time but the time provided for engaging in MVPA was lower than the goal of 60 minutes outlined in the intervention protocol. The average minutes spent in moderate to vigorous physical activity during club sessions was also low (an average of 22 minutes). Therefore, while the fidelity of club time was strong, the use of that time to ensure adequate opportunities for MVPA and engagement in MVPA was lower than intended. Data on minutes spent in management and instruction suggests that the high time spent in group management tasks may have led to low physical activity time. Overall, the data from attendance and fidelity observation indicates that fidelity in terms of attending the club to have opportunities for physical activity, opportunities for physical activity during the session, and actual MVPA engaged in during the club was low, leading to low intervention dosage.

The intervention fidelity and dosage are likely related to the lack of intervention effects on cardiovascular fitness, body composition, and MVPA. One reason they may be related is that



low levels of engaging in MVPA during the club would likely not cause changes in these outcomes measures. The United States Department of Health and Human Resources recommends 60 minutes per day of moderate to vigorous physical activity because that is the amount determined by the department to be necessary for beneficial health outcomes. In comparison, intervention observations suggest an average of 30 minutes during the intervention sessions. It may also be possible that differences in these areas would be found for participants with high rates of session attendance or high rates of MVPA during sessions. Analysis conducted using participants who attended half or more of the sessions did not find significant intervention effects. However, it is possible that because this significantly decreased the sample that the analysis did not have the necessary power to detect effects that were present.

However, this intervention fidelity, dosage, and physical activity outcomes data indicates that the interventions at the two different schools were appropriate to analyze together. The level of fidelity from observations at each school did not differ significantly on any of the measured variables. Therefore, participants who attended the different school's clubs to the same degree likely received similar dosage of the intervention. The number of participants attending each of the four levels of attendance (0, 25, 50, 75% of sessions) also did not differ by school. Therefore, participants at the two sites were receiving similar levels of intervention dosage.

### **Results of Intervention on Executive Functioning**

The first hypothesis that students in the intervention group would improve significantly more on the inhibition and working memory tasks was partially supported in that intervention participants did improve significantly more in the effective function of inhibition, although with a small effect size. On average, participants in the intervention group grew by close to half a standard deviation on the standardized inhibition measure while participants in the control group

grew by about one quarter of a standard deviation. Two previous studies have found significant effects of physical activity interventions on similar flanker tasks (Fisher et al., 2011; Hillman et al., 2014) and Chaddock-Hillman and colleagues (2013) determined that their data indicated an intervention effect. However, Krafft and colleagues (2014) found no intervention effects on flanker performance. Therefore, the current study adds to the evidence that physical activity interventions can improve inhibition performance. In addition, the small effect size seen in this study is similar to that reported by Hillman and colleagues (2014).

The current findings on inhibition effects also extend the previous work in two areas. One important difference between these current findings on inhibition and the previously existing literature is that the participants for this study were significantly older than in previous work. The participants in this study ranged from 10 to 15 years old while previous work has been conducted with participants ages 7 to 11 (Chaddock-Heyman et al., 2013; Hillman et al., 2014; Krafft et al., 2014) or 5 to 6 years old (Fisher et al., 2011). Therefore, together previous research and these results suggest that a physical activity intervention has inhibition benefits in addition to the better established health benefits and extends the age range known to benefit from these results.

However, there are alternative explanations for the results relevant to this first hypothesis. This lack of intervention effect differs from the two previous studies on the effect of a physical activity intervention on working memory which both found intervention effects (Fisher et al., 2011; Kamijo et al., 2011). It is possible that the lack of improvement in working memory skills seen in this study is because increasing physical activity time using an after-school intervention does not truly cause changes in working memory.

While previous work has suggested working memory improvements as a result of increasing physical activity time without decreasing academic time, there are potential

differences in one of the two studies that may mean its results do not truly show changes in working memory. One potential difference is the nature of the working memory measure used. Kamijo and colleagues (2011) used a modified Sternberg task (Sternberg, 1966) where participants were asked to indicate if a single letter had been part of the previously shown set but are not required to tell the examiner all the letters in the set. This differs from the original task where participants first had to repeat the set of letters and then also indicate if a letter had been part of that set. Research indicates that without the recall, like the study here, the task is more similar to visual search tasks than it is to other working memory tasks (Corbin & Marquer, 2013). Therefore, the working memory changes suggested by this previous research may actually be changes in visual recognition or effective visual search strategies. If this is the case, then it might be argued that the work by Kamij and colleagues (2011) does not contradict the findings in this study.

Another possible explanation for the difference in findings between this work and previous work is that visual working memory but not verbal working memory is affected. The CANTAB measures used by Fisher and colleagues (2011) involved two parts. In the first task, similar to the Corsi blocks, squares on a screen changed color and the participant was to touch the boxes in the same order. In the second, participants were expected to find ‘tokens’ hidden behind the boxes until they find all the tokens, which requires visuospatial memory for searched areas, and strategically exploring new locations. Both working memory tasks may have required different aspects of working memory than the task used in this study because they require retention and manipulation of visuospatial information. In addition, unlike for the NIH measure used in this study, the abstract nature of these measures may not allow for or as strongly suggest the use of verbal strategies. Therefore, the effect seen in both studies may have been an effect

specifically because of the visual spatial working memory aspects of the tasks. If this is the explanation for findings it would indicate that physical activity interventions may be a more effective intervention for improving visuospatial working memory than verbal working memory.

An additional reason for a lack of finding of working memory change to consider is that the intervention in this study was not sufficient for causing changes in that executive function area. First, the low dosage due to low attendance and low time in MVPA may have blocked any physical activity induced changes that might have been seen. Kamijo and colleagues (2011) found intervention effects on working memory but had a much higher attendance rate (82% versus 51%). In this previous work, the intervention was also significantly longer than the intervention in this study, at 9 months compared to 17 weeks (about 4 months). Therefore, higher dosage might be necessary for intervention effects on working memory. It is important to note that change in inhibition was seen in the current study even with the lower attendance rates and the shorter intervention length. Therefore, if it is true that no working memory intervention effects were seen because the dosage did not meet the threshold necessary for working memory effects than this would mean that working memory skills require higher physical activity intervention dosages to improve.

Relatedly, this study may not have been sufficient to cause working memory changes is that the intervention in this study did not cause changes in cardiovascular fitness or body composition. The intervention in the work by Kamijo and colleagues (2011) that found working memory effects also caused significantly more improvement in cardiovascular fitness in the intervention group. As discussed in the literature review, these factors may be related to executive function skills. If this is true, changes in these factors may need to be seen for changes particularly in working memory to be seen. Or, these variables may act as indicators that changes

in the proposed biological mechanisms occurred and therefore a lack of change in these areas due to intervention means mechanisms necessary for working memory change did not occur. Kamijo and colleagues (2011) found changes in working memory as a result of an intervention but also found intervention effects on cardiovascular fitness, which supports this hypothesis. Therefore, the lack of impact on cardiovascular fitness and body composition, or on the proposed biological mechanisms that this lack may indicate, may have led to the lack of effect on working memory.

A last possible explanation for the lack of change in working memory may be that physical activity can affect working memory but that it was not seen in this study because of differences in developmental trends in the different executive functions. Research indicates that working memory may continue to develop through young adulthood (Best, Miller, & Jones, 2009). Therefore, it is possible that the participants in the study were not of an age where changes to working memory skills are most amenable to intervention. Both the intervention and control group started at significantly lower levels of working memory performance than intervention and shifting performance. This fits with the idea that working memory was less developed in these participants because of their age. However, participants in the two previous studies that looked at working memory found effects with even younger participants (5 to 6 years and 7 to 9 years) than utilized in the current work. In addition, while an intervention effect was not found, both groups did show a significant improvement working memory over time. It is possible therefore that working memory is not susceptible to large changes due to the more extended period of development or that there may be a bimodal growth curve with two periods of heightened growth or sensitivity, five to nine years old and then again in young adulthood.

The last finding in the area of executive functioning, that no change in shifting was seen, was expected. While one study did identify an intervention effect in shifting, the tool used in that research provided a more detailed analysis (Hillman et al., 2014). The researchers found an effect on accuracy and not reaction time for the more difficult heterogeneous trials. The NIH measure used in this study did not allow for differentiation of heterogeneous and homogeneous trials or for separating out response time and accuracy scores. Therefore, there may be an effect that was obscured by the performance measure available. However, the participants in the current study were also older on average than the participants in the work by Hillman and colleagues (2014). Preliminary research into the development of executive functions suggests that children may reach their mature shifting ability by 10 or 11 years of age (Anderson, 2002; Huizinga, Dolan, & van der Molen, 2006). This differs from inhibition, where Huizinga et al. (2006) found continued improvement in inhibition on a Flankers task until age 15. Because the participants in the current study were in this older range (ages 10 to 15 with mean age of 12) while the participants in Hillman and colleagues (2014) work were not, (7 to 9 years), the lack of finding in shifting performance may also indicate that shifting is not as effected by intervention later in childhood or when approaching adolescence. However, without the ability in this study to look more specifically at shifting performance change in both age groups or to look at response time and accuracy and heterogeneous and homogeneous trials separately, it is not possible to determine which explanation, type of measure or age of participants, is more likely.

### **Results of Intervention on Math Achievement**

The hypothesis that students in the intervention group would improve significantly more on both math achievement measures was partially supported in that intervention participants did improve significantly more in math fluency, although with a small effect size. On average,

participants in the intervention group grew slightly on the measure while participants in the control group decreased their scores slightly. The finding in the current study of changes in math fluency provides preliminary evidence that increasing physical activity can have a significant though medium effect on an important academic outcome. The one previous study specifically examining the effects of an after-school physical activity intervention on math achievement found positive intervention effects on a broad math index that included the two measures in this study (math fluency and applied problems) as well as a third measure, calculation (Davis et al., 2011). Together the results from this study and the work by Davis and colleagues indicates that changes are most likely to be seen in seen in math fluency alone or potentially in math fluency and math calculation skills.

Although hypothesized, no intervention effects were seen in the measure of math problem solving. This may indicate that physical activity, and its associated biological and executive functioning changes, does not have an effect on math problem solving abilities in pre-adolescent children. This lack of effect has multiple possible explanations. First, this lack of effect may be because math problem solving is a more complex, higher level math skill. Improvements in executive functions may have an influence on lower levels math skills, such as making it easier for participants to quickly and accurately recall basic math facts, but not on higher order skills. However, this explanation would not fit with the conceptualization of executive functions as higher order, control functions that are utilized when working towards a more complex goal (Diamond, 2013).

Second, it may be that math problem solving was not changed by the intervention because of the intervention's effects on specific executive function. While changes in shifting or working memory might improve math reasoning, improvements were only seen in the executive

function of inhibition. It may be that if there were changes in shifting and/or working memory as a result of the intervention then there would also have been changes in math problem solving. However, the mediation analysis did not support the hypothesis that a change in inhibition performance mediated the change seen in math fluency. Therefore, it is unclear if the change in math fluency seen was caused by a change in the executive function theorized to be related to that math skill and so other explanations related to the intervention and measures used are more likely.

A last and most likely explanation is that physical activity effects were seen on math fluency but not math problem solving because math fluency is a timed measure while math problem solving is not. One of the proposed effects of increased physical activity on the brain is improved blood oxygenation. In the long term, regular physical activity may enhance cerebral blood flow volume, causing an increase in angiogenesis, the formation of new blood vessels from pre-existing vessels. These new blood vessels result in increased blood flow, which may support the increased metabolic demands (oxygen and glucose) of increased brain function, including in the cerebellum (Black, Isaacs, Anderson, Alcantara, & Greenough, 1990), and cortex (Ding et al., 2006). For example, research with healthy young adults has found that providing them with oxygen improves performance on simple reaction time tasks (Moss, Scholey, & Wesnes, 1998). It is possible that improved blood flow and therefore improved oxygenation and glucose levels in the brain could result in faster cognitive processing speed or faster motor speeds, both of which would affect a speeded motor task like the math fluency without necessarily improving performance on an untimed task like math problem solving.

Lastly, the results of this study failed to show a mediation relation between executive functions and math achievement. It was hypothesized that improvement in math fluency would



be related to shifting and inhibition while any improvements in math problem solving would be related to improvements in working memory and inhibition. Neither hypothesis was supported by the mediation analysis conducted. Performance on tasks measuring all three executive functioning has been previously linked to mathematical performance overall (Bull & Scerif, 2001; St Clair-Thompson & Gathercole, 2006) In addition, associations between working memory and inhibition with math problem solving and of inhibition and shifting with math calculation have been found (Geary, Hamson, & Hoard, 2000; Lan et al., 2011; LeBlanc & Weber-Russell, 1996; Lee et al., 2009; Passolunghi & Siegel, 2001; Swanson & Alloway, 2012). The current study findings on the association between change in executive function and math achievement following an intervention therefore differ from multiple previous correlational studies indicating significant links between executive function and math achievement.

In analyzing the results of the mediation analysis, it is important to remember that different conclusions may be drawn based on how one evaluates the model quality. The first explanation for these findings is that the mediation analysis correctly models the relation between these three variables. If this is true, then the changes in executive function and in math achievement are two, separate, independent outcomes of the physical activity intervention. This would indicate that although the existing work reviewed in the previous paragraph has shown relations between math achievement and executive function, this correlational relation does not indicate a true causal relation where executive function development contributes to math achievement.

However, the lack of mediation effect shown in this model may also be a result of the specification of the model. In other words, the lack of findings could also be attributed to a model that does not correctly model the interaction between executive function and math

achievement. First, the model might need to be a serial model with the executive functions mediating the effect in a certain order rather than the simultaneous model used in this study. Second, the model may need to incorporate moderation as well as mediation in order to accurately reflect the way that changes in executive function effect changes in math achievement. Lastly, it is possible that the use of change scores in the model is too problematic to truly portray the mediation effect. If this is the case, a longitudinal model with more than three times points of mediation may be necessary to accurately capture the relations between these variables. Because these alternative models have yet to be explored, it is important to note that this work is preliminary and conclusions should be postponed until future work in this area is completed.

The lack of support for the mediation effect of working memory on math problem solving is not surprising the lack of intervention effect on working memory. Multiple correlational design studies have shown that working memory uniquely predicts math problem solving performance and that this is a substantial relation (Cragg & Gilmore, 2014). In addition, some working memory intervention studies using cognitive training have found improved math problem solving at follow up months after the intervention (e.g. Holmes, Gathercole, & Dunning, 2009). However, unlike working memory training interventions that have shown effects on math problem solving, the current study may not have significantly improved working memory skills enough in the intervention group. Although there was a significant effect of time on working memory skills, the effect size was smaller than for inhibition. In addition, whereas roughly half of the participants showed any amount of positive change in working memory, roughly three quarters of the inhibition group showed any positive change. Therefore, the most likely reason

that this result may have occurred is that, unlike math fluency and inhibition, the level of change in working memory was not significant enough to cause a change in math problem solving.

A secondary explanation is that, as some researchers in the area argue, the majority of correlational studies showing working memory and math achievement relations have used EF tasks involving numerical stimuli (e.g. digit span). These measures may overestimate the role of EF skills in mathematics compared to non-numerical tasks because of their domain-specificity. In other words, working memory for numbers may be related to math achievement while working memory for other information is not. For example, St. Clair-Thompson and Gathercole (2007) found domain-specific relations between verbal working memory and English achievement and nonverbal working memory and math achievement. However, conflicting findings (Cragg & Gilmore, 2014) of working memory predicting mathematics performance even when non-numeric stimuli were used (e.g. letters or words or remembering block patterns) suggests that this explanation is less likely.

### **Limitations**

It is important however to acknowledge the limitations of the current work in answering the important questions about the potential benefit of physical activity interventions for the academic achievement and executive function of children, such as the inability to rule out certain alternative explanations for the findings or to determine the specific components of the intervention that are the essential therapeutic elements. An examination of these limitations below will help to ensure that the findings from this study are utilized appropriately and will also be used to guide a discussion on future research.

One area of limitation in the current study is the specificity of the measures used, as the National Institute of Health's Flanker and Dimensional Card Sort tasks provide scores that are less specific than the tasks used in other studies. First, these measures provides one combined score rather than one score for the participants' accuracy and one score for their response time. Therefore, it is difficult to directly compare findings from this study to findings by Hillman and colleagues (2014) and Fisher and colleagues (2011). The Dimensional Card Sort task used in this study to assess shifting also does not provide separate scores for heterogeneous (the rule switches during the testing block) and homogeneous trials (the rule stays the same during the testing block). Therefore, it is not possible to directly compare with Hillman and colleagues (2014) findings of an effect on accuracy for heterogeneous trials but not for accuracy on homogeneous trials. Together, these limits of scoring create limitations for comparing these results to others and therefore for understanding the details of these effects.

Second, there are multiple study design and measure limitations that make it difficult to rule out alternative explanations. The current study measures do not allow for an examination of an alternative explanation of the results based on participant mental health and motivation. It is possible that individuals in the intervention group had higher levels of depression and that the physical activity club improved achievement and executive function performance because of improved depressive symptoms in this group. The majority of studies that have examined the efficacy of exercise to reduce symptoms of depression and anxiety in children have described a positive benefit associated with exercise involvement, particularly with structured involvement (Strong et al, 2005). Studies have shown that adolescents with depression show lower grade performance and parent ratings of school functioning (Jaycox, 2009). Depression is often associated with an inability to concentrate and intrusive ruminative thoughts, which are likely to

reduce available cognitive resources for performance on academic and executive function measures. However, because no measures of depression or anxiety were included in this study, it is not possible to determine if this alternative explanation explains the findings in this study.

The low intervention fidelity and lack of changes in cardiovascular fitness and body composition limit the conclusions that can be drawn from negative findings. Low attendance to the intervention as well as low amounts of time in moderate to vigorous physical activity during the intervention negatively affect the dosage of physical activity. In addition, there were no changes in cardiovascular fitness or body composition as a result of this intervention. It is likely that for changes to be seen in more advanced executive function skills (e.g. working memory) and academic skills (e.g. math problem solving) that more intensive intervention would be necessary. As discussed previously, there are multiple proposed biological mechanisms by which increased physical activity may cause functional and structural changes in the brain. However, the dosage for these changes may be much higher than seen in this intervention. For example, while research in dyslexia has shown that intervention can change in activation patterns in the brain, this happens after intensive intervention, such as fifty minutes daily for one year (Shaywitz et al, 2004). Therefore, the lack of findings for hypothesized changes in shifting and math problem solving cannot conclusively be determined to mean that physical activity intervention does not affect these outcomes. The alternate explanation that physical activity interventions with higher dosage could influence shifting, working memory, and math problem solving cannot be refuted with the data from this study.

Second, the design of the study limits what conclusions can be drawn about mediating factors and essential elements of the intervention. The control group did not receive intervention of any kind. The investigation of a proposed mediator of treatment is best conducted when

mediators are investigated with a control condition that allows for ruling out the possibility that what appears to mediate change is simply a general effect of receiving treatment (Murphy, Cooper, Hollon, & Fairburn, 2009). Without having an active control group to compare the physical activity intervention group to it is not possible to determine clearly what aspects of the intervention are responsible for the differential changes in math fluency and inhibition. Second, because assessment was completed during the same time points as assessment for the Girls on the Move Intervention to minimize disruption to the school and avoid removing participants from the after-school club, measures of executive function were not able to be taken at midpoints between pre-test and the end of the intervention. A crucial piece of assessment design for mediator analysis is the timing of the measurements (Kazdin, 2007). Measuring both the potential mediator and the outcome variables at frequent intervals is helpful for determining if the mediator changes prior to the change in the outcome variable as would be expected. In this study, a the potential mediator (executive function) was not measured at any time point before post-test math achievement measures were collected. Therefore, determinations cannot be made from this study about if changes in inhibition happened prior to the significant change in math fluency performance seen at post-test. Together, these study design limitations mean that the mechanisms of change in this study can be hypothesized but not directly tested.

### **Future Research**

Future studies should aim to provide specific additional data that helps to better understand the findings of the current study related to the interventions effects and differential effects. First, a measure of depression should be included in future work in order to determine any differences at baseline between the intervention and control groups' depressive symptomology and to allow for an analysis of change in depressive symptoms as a possible

causal mechanism for changes in executive function and math achievement. Second, the working memory measured used in this study gave participants both visual and verbal information to store and manipulate which made it impossible to determine differential effects on verbal and visual working memory which have been suggested by previous research. Future research should therefore include separate measures of verbal and visuospatial working memory. Third, it is possible that intervention effects were found on only the timed math measure may indicate that the effect is on speed of processing or motor speed and not specifically on math fact fluency. Therefore, research that includes a measure of processing speed, such as a cancellation or coding measure, and a measure of motor speed, would be helpful for determining if the effect of the intervention on math fluency is specific to math knowledge or is due to change in processing or motor speed. Last, the current study examined only one specific academic area, math. It will be very important for similar research to be conducted that examines the effect of physical activity on multiple areas of reading and writing and whether any executive functions act as a mediator on those effects if they are seen. The addition of these specific measures to similar future research will help to better understand the results from this current study.

Now that multiple randomized control trial intervention studies have shown effects of physical activity interventions on executive function, the next step for research in this area will be the use of multi-level longitudinal models. In the current, study, only pre-test and post-test measures of both executive function and math achievement were completed. Therefore, no measures of long term follow up were completed. Without any longer term follow up, it is not possible to determine if changes in inhibition and math fluency were maintained and if so, how long the effects would be maintained. While the current study provides evidence for an

intervention effect of physical activity on inhibition and math fluency, it is important for future research to analyze the long-term outcomes by completing assessments at multiple time points.

Empirical investigations completed in schools, like the current study, are inherently hierarchical because even though results are measured at the level of the student, these students are nested in classrooms, in grade levels, and in schools. Therefore, individuals that are considered independent participants are actually hierarchically ordered groups that experience their educations and the intervention in the same contexts. Many researchers in the area of education argue that children's learning is strongly influenced by the educational context in which it occurs (Lee, 2000). When children attend school, they experience education in small groups, such as grade levels, classrooms, or differentiated groups within classrooms. In the current study, data was analyzed at the level where the intervention was thought to occur (i.e. the student level). However, because the data is in multilevel, misestimated standard errors could have occurred because individual cases were treated as independent case when in fact they were not (Raudenbush & Bryk, 1986). Students' achievement in the same district and school share at least some dependence with other schoolmates because they share many educational experiences and students are assigned to both districts and schools by location rather than randomly. If the assumption made, that relations between student personal characteristics and outcomes are similar regardless of school, is incorrect then the effect of school is being underestimated. Therefore, use of hierarchical linear models in future analysis of this question will be important for understanding the true contribution of the effect of school on this question.

In addition, the Girls on the Move after-school club included multiple components in addition to the physical activity time including time with other adults and children, nutritious snacks, diverse physical activities, and motivational interviewing. First, in the present study,



children in the intervention group had social interaction with club personnel and each other during the physical activity. Next, children in the intervention group received a healthy snack. Research shows that dietary factors influence specific molecular systems and mechanisms that maintain mental function, including synaptic plasticity (Gomez-Pinilla, 2010). This may be particularly important as the intervention and control group did differ significantly in participation in the Free and Reduced Lunch program, which may mean that the two groups differed in their access to food, particularly healthy food. Third, although the physical activity program was designed to enhance cardiorespiratory fitness, the activities sometimes also involved other factors such as following rules for active games or working together in a group, that might also influence children's cognitive function (Best, 2010). Unfortunately, with the data available it is not possible to rule out these other possible explanations for an intervention effect or to conclusively determine that engaging in physical activity is the cause of the intervention effect seen on inhibition and academic performance. Therefore, study designs aimed at better analysis of the essential components of the interventions provided in this and future studies are now necessary.

Relatedly, now that the selective effects of physical activity interventions on executive functions and math achievement have begun to be understood, it will be important to take the next step to understand the dosage necessary to have these effects. Specifically, the collection of individual participant data on physical activity level during the physical activity will be an important component of future research. Analysis of how level of physical activity during session relates to outcomes and if there are certain thresholds of physical activity necessary to see the effects is an important next step in this area of research as it will allow for practical recommendations for educators.

In summary, understanding when and why generally evidence-based interventions work or do not work is a difficult task for psychologists, as it requires thoughtful and complex designs and methods of analysis (Kazdin, 2007; Kraemer, Wilson, Fairburn, & Agras, 2002). While this area of research is interdisciplinary in nature, it is no different in this regard. Future research utilizing an active control, assessment of specific intervention components and individual dosage, and a multi-level longitudinal mediation analysis of the relation between executive function and academic achievement after a physical activity intervention would help to answer this call in regards to physical activity interventions for these outcomes.

### **Implications**

The findings of this study have important implications for educators and education researchers, although the limitations of these must be acknowledged. The findings that increasing physical activity of students can improve not only the executive function of inhibition but also their math fluency performance provides further evidence that educators can add physical activity interventions to their list of tools. In addition, the fact that these intervention effects occurred despite the low intervention fidelity suggests that the findings from this study could be reproduced without need for educators to provide prohibitively high levels of support in terms of student attendance or instructor ability to maximize time in moderate to vigorous physical activity.

However, implications for educators are limited by two aspects of the study methods. First, while an after-school intervention model may work for some schools, it does not directly inform educators who may be struggling with decisions about how much instructional time versus recess and physical education time to provide to their students during the day. Specifically, this study involved an after-school physical activity club and therefore did not take

away from instructional time in the school day. Another area of limitation for the implications of the study for educators is that the findings may not generalize to other populations. Participants were all female, between the ages of 10 and 15, and attended school in districts that primarily serve low-income, racial minority student populations. In addition, the participants were specifically chosen because they engaged in low levels of physical activity and were not otherwise occupied with after-school activities. Therefore, these findings may not be generalizable to populations of high socioeconomic status or in schools where there are significant other enrichment and exercise based activities available and may not apply to older adolescent students or to males.

Together with the results from the previous two research questions, the information from the analysis of mediation provides guidance for the continuing discussions of existing findings on the effect of physical activity interventions on executive functions. As reviewed previously, there are many correlational and longitudinal studies linking executive functions to math achievement. The evidence has suggested that both working memory and inhibition contribute to math achievement beyond the contributions of intelligence. Because of this data, researchers in the area of physical activity and executive functions frequently identify improving academic performance as a reason for the importance of their findings on executive function. They indicate that because, in their view, we know that executive functions are necessary for academic achievement then any intervention improvements in executive function are important for educators. However, analysis of results in change in math and executive function as a result of the same intervention is something done by few studies (Davis et al., 2011; Donnelly & Lambourne, 2011). This is one of few studies that directly measures both executive function and academic achievement after a physical activity intervention and then looks at the relation

between the change in these areas. The interest shown by neuropsychology and education researchers in identifying important real world outcomes of research is understandable and the changes in both math fluency and inhibition after a physical activity intervention does confirm previous findings that a physical activity intervention can improve both areas. However, the lack of evidence for executive function as a mechanism by which physical activity affects academic achievement indicates that these claims should not yet be made.

### **Concluding Remarks**

The current study contributes important information for clinicians and researchers despite limitations. First, this study shows that a physical activity intervention that adds physical activity into the day of school-aged children improves their inhibition and math fluency skills. The success of this intervention in these areas in a school serving a primarily low-income, minority student population suggests that this may be an effective intervention route for schools looking to improve their students' physical health as well as academic achievement. The lack of evidence for executive function's mediation effect in this preliminary work suggests that researchers in the area of physical activity and executive functions need to continue to examine this potential mediator prior to utilizing the connection as a rationale for the importance of their work for parents and educators.

## APPENDICES

*Appendix A*  
*Parent and Student Consent Forms*

**Study Title:** Academics and Exercise

**Researchers:** Katelin daCruz, M.A. Graduate Student in School Psychology  
Jodene Goldenring Fine, Ph.D., Assistant Professor of School Psychology

**Purpose of the Research:** The goal of this study is to see if participating in a physical activity club can help girls with their performance in school.

**Participation:** This study is an *optional* additional study. Girls who participate in the “Girls on the Move Intervention” have the opportunity to also participate in this study. If the study receives more participation will be decided on a first come, first served basis.

**Description of the Project:** If your daughter is selected:

- She will be asked to do some short academic tasks (e.g. reading words, math facts) twice this year
- She will be asked to do some short tasks that tell us about her attention and memory twice this year
- Her grades and attendance will be reviewed by the researchers during the school year.
- Information about your daughter from the Girls on the Move Intervention (e.g. age, height, and amount of physical activity) will also be reviewed.

Some tasks will be on the computer. Some tasks will be done with the researchers. The total amount of time to complete these will range from 20 to 30 minutes. The assessments will be done at her school and take place during the school day. The standardized scores from the assessment will be provided to you in a brief score report if you request it.

**Your Rights to Participate, Say No, or Withdraw**

Your decision to allow your child to participate in this research study would be greatly appreciated. However, your child’s participation in the study is entirely voluntary. Not participating in this additional study will not affect participation in the Girls on the Move Intervention. You and your child have the right to say no. You or your child may change your minds at any time and withdraw from the study. Whether or not you or your child chooses to participate will have no effect on your child’s grade at school and no information gathered during the study will be shared with your child’s school or teacher.

**Potential Benefits**

Participating in one-on-one testing could be a positive experience for your child. Other potential benefits of participating in this study include helping us to better understand how physical activity effects academic performance.

**Potential Risks**

There is minimal risk in participating in the study. Your child may experience some anxiety if she or he tends to get anxious during testing. However, the researchers will address any of your child’s concerns or questions before, during, or after the testing. Your child will be allowed to take breaks if needed, and will be provided with encouragement and feedback when appropriate.

**Costs and Compensation for Being in the Study**

It does not cost anything to participate in this study. Your child will earn 5 dollars for the first testing session and 10 dollars for second testing session, because it will take more time.

**Privacy and Confidentiality**

The data from this research study will be kept confidential. Information about you will also be kept confidential to the maximum extent allowable by law. All identifiable and de-identified data will be stored in a locked file cabinet and password-protected computer. Data from the child's assessment will be entered into a database by his/her subject number and the information linking the study data to individuals' identities will be kept separately and securely. At no point will the child's testing performance be disclosed to teachers or others. A logbook linking your child's name and research number will be kept in a locked file cabinet and electronically on a password-protected computer. Only the researchers and the MSU Institutional Review Board (IRB), which is responsible for oversight of the safety and rights of research participants, will have access to the data. In the event that results of this research project is presented at a professional conference or published journal article, the child's identity will be disguised and other personally identifiable information will not be disclosed.. It will be kept for 5 years following the completion of the study and then will be destroyed.

**Contact Information for Questions or Concerns**

If you have concerns or questions about this study, such as scientific issues, how to do any part of it, or to report an injury, please contact the researchers:

**Katelin daCruz**

Graduate Student in School Psychology  
344 Erickson Hall  
East Lansing MI 48824  
(616) 520-8466  
dacruzka@msu.edu

**Jodene Fine, Ph.D**

Assistant Professor of School Psychology  
440 Erickson Hall  
East Lansing MI 48824  
(517) 884-0443  
finej@msu.edu

If you have questions or concerns about your role and rights as a research participant, would like to obtain information or offer input, or would like to register a complaint about this study, you may contact the following, anonymously if you wish:

Michigan State University's Human Research Protection Program  
202 Olds Hall, 408 W Circle Drive, Michigan State University, East Lansing, MI 48824  
Phone: 517-355-2180, Fax 517-432-4503

**Title of the Study:** Academics and Exercise

**About the Study:** The goal of this study is to see if participating in a physical activity club can help girls with their performance in school.

- You will do some short academic tasks (e.g. reading words, math facts) twice this year
- You will do some short tasks that tell us about your attention and memory twice this year
- Your grades and attendance will be reviewed by the researchers during the school year

The total amount of time to complete these will range from 20 to 30 minutes. Some tasks will be on the computer. Some tasks will be done with the researchers. The assessments will be done at your school and take place during the school day.

**Who Can Participate:** Girls who participate in the “Girls on the Move Intervention” can do this study too. If you agree to participate in that study, you can also agree to participate in this study if you want to.

**What if I don’t want to do this study:** You can still participate in the “Girls on the Move Intervention”! You choose if you want to be in the study or not. If you decide to quit the study after we start, all you have to do is tell the person in charge.

**Your Privacy:** Only the people in charge of the study will know what you say and do in the study.

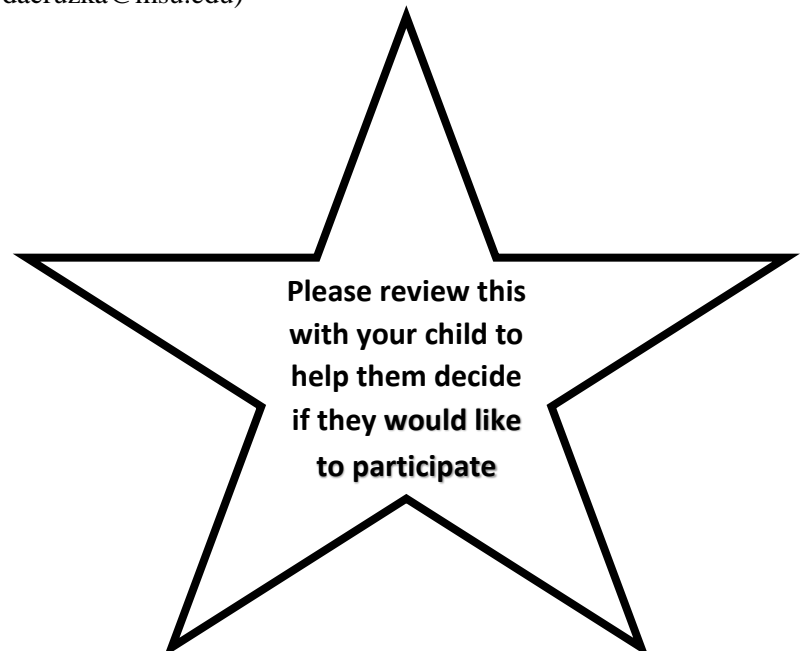
**Potential Benefits:** You will be helping us learn about how physical activity can help other kids your age.

**Potential Risks:** You could feel nervous or anxious during the testing.

**Costs:** Participation is FREE

**Reward:** \$5 for the first session and \$10 for the second session (it takes a little longer)

**Contact Information:** If you have concerns or questions about this study, please contact the researcher Katelin daCruz ((616) 520-8466 or by email: [dacruzka@msu.edu](mailto:dacruzka@msu.edu))





**Please return this form *signed***  
**if you wish to enroll your daughter in the Academics and Exercise study**

**1. Parent/Guardian Consent**

I have received a copy of the consent form and I agree to allow my child to participate in this research study.

Parent/Guardian Printed Name: \_\_\_\_\_

\_\_\_\_\_  
Signature of Parent or Guardian

Today's Date \_\_\_\_\_

Child's Name \_\_\_\_\_ Child's Birth Date \_\_\_\_\_

**2. Student Assent**

I have read the consent form about the "Academics and Exercise" study, or an adult has read it to me. I am signing below because I want to participate in this research project.

STUDENT Printed Name: \_\_\_\_\_

STUDENT Signature: \_\_\_\_\_ Date: \_\_\_\_\_

*Appendix B*

*Timeline for completion of study*

<b>Activities</b>	<b>Sept 2014</b>	<b>Oct. 2014</b>	<b>Nov. 2014</b>	<b>Dec. 2014</b>	<b>Jan 2015</b>	<b>Feb. 2015</b>	<b>Mar. 2015</b>	<b>Apr. 2015</b>
Recruitment of girls from 4 different schools	X	X						
Baseline data collection		X	X					
After-School PA Club Begins at Intervention Schools		X	X	X				
Scoring of baseline data.					X	X		
After-school PA Club ends. Post-intervention data collection.							X	X

Appendix C

Variable Correlations

Table 21. Variable Correlations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Age	-																		
2. Pre Body Fat	.068	-																	
3. Pre-test BMI	-.011	<b>.882</b>	-																
4. PreVO2	-.164	<b>-.925</b>	<b>-.894</b>	-															
5. Pre MVPA	-.032	-.053	-.075	.106	-														
6. Post BodyFat	.106	<b>.904</b>	<b>.906</b>	<b>-.952</b>	-.088	-													
7. Post BMI	.023	<b>.866</b>	<b>.969</b>	<b>-.897</b>	-.043	<b>.949</b>	-												
8. PostVO2	-.151	<b>-.894</b>	<b>-.856</b>	<b>.968</b>	.078	<b>-.967</b>	<b>-.901</b>	-											
9. Post MVPA	-.186	-.137	-.106	.160	.525	-.133	-.063	.113	-										
10. Pre WJMF	-.139	-.241	-.205	.258	.030	-.232	-.216	.250	-.053	-									
11. Pre WJAP	<b>-.303</b>	-.199	-.136	.218	-.065	-.203	-.177	.217	-.058	<b>.491</b>	-								
12. Pre NIH DCCS	-.113	-.178	-.138	.157	-.014	-.153	-.162	.175	.043	<b>.323</b>	<b>.294</b>	-							
13. Pre NIH FIT	-.168	-.057	-.030	.076	.082	-.057	-.034	.063	.040	<b>.308</b>	<b>.322</b>	<b>.367</b>	-						
14. Pre NIH LS	<b>-.302</b>	-.083	-.106	.139	.102	-.096	-.108	.116	.068	.190	<b>.496</b>	<b>.209</b>	.256	-					
15. Post WJMF	-.127	-.252	-.230	.270	-.039	-.246	-.240	.272	-.087	<b>.864</b>	<b>.518</b>	<b>.354</b>	<b>.323</b>	.179	-				
16. Post WJAP	<b>-.360</b>	-.202	-.122	.210	-.076	-.181	-.156	.196	-.080	<b>.474</b>	<b>.796</b>	<b>.311</b>	<b>.359</b>	<b>.451</b>	<b>.531</b>	-			
17. Post NIH DCCS	-.095	-.169	-.133	.127	.047	-.174	-.184	.170	.058	.169	.199	<b>.619</b>	<b>.320</b>	.138	.251	.204	-		
18. Post NIH FIT	-.209	-.197	-.141	.207	-.016	-.186	-.168	.212	-.042	<b>.293</b>	<b>.373</b>	<b>.304</b>	<b>.649</b>	.257	.336	<b>.408</b>	<b>.394</b>	-	
19. Post NIH LS	-.133	-.049	-.033	.085	.158	-.061	-.029	.071	.028	.160	<b>.473</b>	.165	.249	<b>.459</b>	.185	<b>.418</b>	.178	.243	-
20. Club Attend.	.006	-.087	-.037	.057	-.079	-.049	-.038	.048	-.018	-.038	.038	-.098	-.100	-.002	.037	.069	.028	.072	-.002

Correlations between variables. Bolded variables have p values less than .001

## REFERENCES

## REFERENCES

- Ahamed, Y., Macdonald, H., Reed, K., Naylor, P.-J., Liu-Ambrose, T., & McKay, H. (2007). School-Based Physical Activity Does Not Compromise Children's Academic Performance. *Medicine & Science in Sports & Exercise*, *39*(2), 371–376. <http://doi.org/10.1249/01.mss.0000241654.45500.8e>
- Altman, J. (1962). Are new neurons formed in the brains of adult mammals? *Science (New York, N.Y.)*, *135*(3509), 1127–1128.
- Amis, J. M., Wright, P. M., Dyson, B., Vardaman, J. M., & Ferry, H. (2012). Implementing Childhood Obesity Policy in a New Educational Environment: The Cases of Mississippi and Tennessee. *American Journal of Public Health*, *102*(7), 1406–13.
- Anderson, P. (2002). Assessment and Development of Executive Function (EF) During Childhood. *Child Neuropsychology*, *8*(2), 71–82. <https://doi.org/10.1076/chin.8.2.71.8724>
- Andersson, U., & Lyxell, B. (2007). Working memory deficit in children with mathematical difficulties: A general or specific deficit? *Journal of Experimental Child Psychology*, *96*(3), 197–228. <http://doi.org/10.1016/j.jecp.2006.10.001>
- Banich, M. T. (2009). Executive function: The search for an integrated account. *Current Directions in Psychological Science*, *18*(2), 89–94. <http://doi.org/10.1111/j.1467-8721.2009.01615.x>
- Bassett, D. R., Fitzhugh, E. C., Heath, G. W., Erwin, P. C., Frederick, G. M., Wolff, D. L., ... Stout, A. B. (2013). Estimated Energy Expenditures for School-Based Policies and Active Living. *American Journal of Preventive Medicine*, *44*(2), 108–113.
- Beets, M. W., Beighle, A., Erwin, H. E., & Huberty, J. L. (2009). After-school program impact on physical activity and fitness: a meta-analysis. *American Journal of Preventive Medicine*, *36*(6), 527–537. <http://doi.org/10.1016/j.amepre.2009.01.033>
- Beets, M., Wallner, M., & Beighle, A. (2010). Defining standards and policies for promoting physical activity in afterschool programs. *Journal of School Health*, *80*(8), 411–417. <http://doi.org/10.1111/j.1746-1561.2010.00521.x>
- Best, J. R. (2010). Effects of Physical Activity on Children's Executive Function: Contributions of Experimental Research on Aerobic Exercise. *Developmental Review : DR*, *30*(4), 331–551.
- Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between Executive Function and Academic Achievement from Ages 5 to 17 in a Large, Representative National Sample. *Learning and Individual Differences*, *21*(4), 327–336. <http://doi.org/10.1016/j.lindif.2011.01.007>

- Black, J. E., Isaacs, K. R., Anderson, B. J., Alcantara, A. A., & Greenough, W. T. (1990). Learning causes synaptogenesis, whereas motor activity causes angiogenesis, in cerebellar cortex of adult rats. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(14), 5568–5572.
- Boreham, C., & Riddoch, C. (2001). The physical activity, fitness and health of children. *Journal of Sports Sciences*, *19*(12), 915–929. <http://doi.org/10.1080/026404101317108426>
- Brownson, R. C., Boehmer, T. K., & Luke, D. A. (2005). Declining rates of physical activity in the United States: what are the contributors? *Annual Review of Public Health*, *26*, 421–443. <http://doi.org/10.1146/annurev.publhealth.26.021304.144437>
- Buck, S. M., Hillman, C. H., & Castelli, D. M. (2008). The relation of aerobic fitness to stroop task performance in preadolescent children. *Medicine and Science in Sports and Exercise*, *40*(1), 166–172. <http://doi.org/10.1249/mss.0b013e318159b035>
- Bull, R., Johnston, R. S., & Roy, J. A. (1999). Exploring the roles of the visual-spatial sketch pad and central executive in children's arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology*, *15*(3), 421–442. <http://doi.org/10.1080/87565649909540759>
- Burkhalter, T. M., & Hillman, C. H. (2011). A Narrative Review of Physical Activity, Nutrition, and Obesity to Cognition and Scholastic Performance Across the Human Lifespan. *Advances in Nutrition: An International Review Journal*, *2*(2), 201S–206S. <http://doi.org/10.3945/an.111.000331>
- Carro, E., Nuñez, A., Busiguina, S., & Torres-Aleman, I. (2000). Circulating Insulin-Like Growth Factor I Mediates Effects of Exercise on the Brain. *The Journal of Neuroscience*, *20*(8), 2926–2933.
- Castelli, D. M., Hillman, C. H., Buck, S. M., & Erwin, H. E. (2007). Physical fitness and academic achievement in third- and fifth-grade students. *Journal of Sport & Exercise Psychology*, *29*(2), 239–252.
- Censabella, S., & Noël, M.-P. (2008). The inhibition capacities of children with mathematical disabilities. *Child Neuropsychology*, *14*(1), 1–20. <http://doi.org/10.1080/09297040601052318>
- Center for Educational Policy. (2007). *Choices, changes, and challenges: curriculum and instruction in the NCLB era*. Washington, DC: Center for Education Policy.
- Centers for Disease Control. (2012). *Youth Risk Behavior Surveillance* (Morb Mortal Wkly Rep No. SS-4) (pp. 1–28).
- Centers for Disease Control. (2014). *Youth Risk Behavior Surveillance- United States, 2013* (Morbidity and Mortality Weekly Report No. Vol. 63 No. 4).

- Chaddock, L., Erickson, K. I., Prakash, R. S., Kim, J. S., Voss, M. W., Vanpatter, M., ... Kramer, A. F. (2010). A neuroimaging investigation of the association between aerobic fitness, hippocampal volume, and memory performance in preadolescent children. *Brain Research, 1358*, 172–183. <http://doi.org/10.1016/j.brainres.2010.08.049>
- Chaddock-Heyman, L., Erickson, K. I., Voss, M. W., Knecht, A. M., Pontifex, M. B., Castelli, D. M., ... Kramer, A. F. (2013). The effects of physical activity on functional MRI activation associated with cognitive control in children: a randomized controlled intervention. *Frontiers in Human Neuroscience, 7*. <http://doi.org/10.3389/fnhum.2013.00072>
- Cizek, G. J. (2003). Review of the Woodcock-Johnson III. In B. S. Plake, J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook*. Lincoln, NE: Buros Institute of Mental Measurements.
- Coe, D. P., Pivarnik, J. M., Womack, C. J., Reeves, M. J., & Malina, R. M. (2006). Effect of physical education and activity levels on academic achievement in children. *Medicine and Science in Sports and Exercise, 38*(8), 1515–1519. <http://doi.org/10.1249/01.mss.0000227537.13175.1b>
- Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science, 14*(2), 125–130. <http://doi.org/10.1111/1467-9280.t01-1-01430>
- Cooke, S. F., & Bliss, T. V. P. (2006). Plasticity in the human central nervous system. *Brain: A Journal of Neurology, 129*(Pt 7), 1659–1673. <http://doi.org/10.1093/brain/awl082>
- Corbin, L., & Marquer, J. (2013). Is Sternberg's memory scanning task really a short-term memory task? *Swiss Journal of Psychology, 72*(4), 181–196. <http://doi.org/10.1024/1421-0185/a000112>
- Cotman, C. W., & Berchtold, N. C. (2002). Exercise: a behavioral intervention to enhance brain health and plasticity. *Trends in Neurosciences, 25*(6), 295–301. [http://doi.org/10.1016/S0166-2236\(02\)02143-4](http://doi.org/10.1016/S0166-2236(02)02143-4)
- Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education, 3*(2), 63–68. <http://doi.org/10.1016/j.tine.2013.12.001>
- Cserjési, R., Molnár, D., Luminet, O., & Lénárd, L. (2007). Is there any relationship between obesity and mental flexibility in children? *Appetite, 49*(3), 675–678. <http://doi.org/10.1016/j.appet.2007.04.001>
- Daley, A. J., & Ryan, J. (2000). Academic Performance And Participation In Physical Activity By Secondary School Adolescents. *Perceptual and Motor Skills, 91*, 531–534.

- Datar, A., & Sturm, R. (2006). Childhood overweight and elementary school outcomes. *International Journal of Obesity (2005)*, *30*(9), 1449–1460. <http://doi.org/10.1038/sj.ijo.0803311>
- David, C. V. (2012). Working memory deficits in Math learning difficulties: a meta-analysis. *International Journal of Developmental Disabilities*, *58*(2), 67–84. <http://doi.org/10.1179/2047387711Y.0000000007>
- Davis, C. L., Tomporowski, P. D., Boyle, C. A., Waller, J. L., Miller, P. H., Naglieri, J. A., & Gregoski, M. (2007). Effects of aerobic exercise on overweight children's cognitive functioning: a randomized controlled trial. *Research Quarterly for Exercise and Sport*, *78*(5), 510–519.
- Davis, C. L., Tomporowski, P. D., McDowell, J. E., Austin, B. P., Miller, P. H., Yanasak, N. E., ... Naglieri, J. A. (2011). Exercise Improves Executive Function and Achievement and Alters Brain Activation in Overweight Children: A Randomized Controlled Trial. *Health Psychology*, *30*(1), 91–98. <http://doi.org/10.1037/a0021766>
- Department of Health and Human Services and Department of Education. (2000). *Promoting better health for young people through physical activity and sports* (A Report to the President from the Secretary of Health and Human Services and the Secretary of Education). Silver Spring MD.
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. <http://doi.org/10.1146/annurev-psych-113011-143750>
- Ding, Y.-H., Li, J., Zhou, Y., Rafols, J. A., Clark, J. C., & Ding, Y. (2006). Cerebral angiogenesis and expression of angiogenic factors in aging rats after exercise. *Current Neurovascular Research*, *3*(1), 15–23.
- Donnelly, J. E., Greene, J. L., Gibson, C. A., Smith, B. K., Washburn, R. A., Sullivan, D. K., ... Williams, S. L. (2009). Physical Activity Across the Curriculum (PAAC): a randomized controlled trial to promote physical activity and diminish overweight and obesity in elementary school children. *Preventive Medicine*, *49*(4), 336–341. <http://doi.org/10.1016/j.ypmed.2009.07.022>
- Dwyer, T., Sallis, J. F., Blizzard, C. L., Lazarus, R., & Dean, K. (2001). Relation of Academic Performance to Physical Activity and Fitness in Children. *Pediatric Exercise Science*, *13*(3), 225–237.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. <http://doi.org/10.3758/BF03203267>
- Etnier, J. L., & Chang, Y.-K. (2009). The effect of physical activity on executive function: a brief commentary on definitions, measurement issues, and the current state of the literature. *Journal of Sport & Exercise Psychology*, *31*(4), 469–483.



- Fairclough, S. J., & Stratton, G. (2005). Physical Activity Levels in Middle and High School Physical Education: A Review. *Pediatric Exercise Science*, (17), 217–236.
- Fairclough, S. J., & Stratton, G. (2006). A Review of Physical Activity Levels during Elementary School Physical Education. *Journal of Teaching in Physical Education*, 25(2), 239–257.
- Fedewa, A. L., & Soyeon, A. (2011). The effects of physical activity and physical fitness on children’s achievement and cognitive outcomes: A meta-analysis. *Research Quarterly for Exercise and Sport*, 82(3), 421–532.
- Fisher, A., Boyle, J. M., Paton, J. Y., Tomporowski, P., Watson, C., McColl, J. H., & Reilly, J. J. (2011a). Effects of a physical education intervention on cognitive function in young children: randomized controlled pilot study. *BMC Pediatrics*, 11, 97.  
<http://doi.org/10.1186/1471-2431-11-97>
- Fox, C. K., Barr-Anderson, D., Neumark-Sztainer, D., & Wall, M. (2010). Physical Activity and Sports Team Participation: Associations With Academic Outcomes in Middle School and High School Students. *Journal of School Health*, 80(1), 31–37.  
<http://doi.org/10.1111/j.1746-1561.2009.00454.x>
- Frieden, T., Jaffe, H., Cono, J., Richards, C., & Iademarco, M. (2014). *Youth Risk Behavior Surveillance—United States. MMWR 2014;63(SS-4)* (Morbidity and Mortality Weekly Report). Atlanta, Georgia: Centers for Disease Control and Prevention.
- Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and Arithmetical Cognition: A Longitudinal Study of Process and Concept Deficits in Children with Learning Disability. *Journal of Experimental Child Psychology*, 77(3), 236–263.  
<http://doi.org/10.1006/jecp.2000.2561>
- Gershon, R. C., Wagster, M. V., Hendrie, H. C., Fox, N. A., Cook, K. F., & Nowinski, C. J. (2013). NIH toolbox for assessment of neurological and behavioral function. *Neurology*, 80(11 Suppl 3), S2-6. <https://doi.org/10.1212/WNL.0b013e3182872e5f>
- Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., ... Inglis, M. (2013). Individual Differences in Inhibitory Control, Not Non-Verbal Number Acuity, Correlate with Mathematics Achievement. *PLOS ONE*, 8(6), e67374.  
<http://doi.org/10.1371/journal.pone.0067374>
- Gómez-Pinilla, F. (2008). Brain foods: the effects of nutrients on brain function. *Nature Reviews. Neuroscience*, 9(7), 568–578. <https://doi.org/10.1038/nrn2421>
- Hall, C. D., Smith, A. L., & Keele, S. W. (2001). The impact of aerobic activity on cognitive function in older adults: A new synthesis based on the concept of executive control. *European Journal of Cognitive Psychology*, 13(1–2), 279–300.  
<http://doi.org/10.1080/09541440126012>

- Herting, M. M., & Nagel, B. J. (2012). Aerobic fitness relates to learning on a virtual morris water task and hippocampal volume in adolescents. *Behavioural Brain Research*, *233*(2), 517–525. <http://doi.org/10.1016/j.bbr.2012.05.012>
- Hillman, C. H., Buck, S. M., Themanson, J. R., Pontifex, M. B., & Castelli, D. M. (2009). Aerobic Fitness and Cognitive Development: Event-Related Brain Potential and Task Performance Indices of Executive Control in Preadolescent Children. *Developmental Psychology*, *45*(1), 114.
- Hillman, C. H., Pontifex, M. B., Castelli, D. M., Khan, N. A., Raine, L. B., Scudder, M. R., ... Kamijo, K. (2014). Effects of the FITKids Randomized Controlled Trial on Executive Control and Brain Function. *Pediatrics*, peds.2013-3219. <http://doi.org/10.1542/peds.2013-3219>
- Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, *12*(4), F9–F15. <http://doi.org/10.1111/j.1467-7687.2009.00848.x>
- Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017–2036. <http://doi.org/10.1016/j.neuropsychologia.2006.01.010>
- Institute of Medicine. (2013). *Educating the Student Body: Taking Physical Activity and Physical Education to School*. Washington, D.C.: National Academies Press.
- Institute of Medicine (U.S.), Committee on Fitness Measures and Health Outcomes in Youth, Pate, R. R., Oria, M., & Pillsbury, L. (2012). *Fitness measures and health outcomes in youth*. Washington, D.C.: National Academies Press.
- Janak, J. C., Gabriel, K. P., Oluyomi, A. O., Pérez, A., Kohl, H. W., & Kelder, S. H. (2014). The Association Between Physical Fitness and Academic Achievement in Texas State House Legislative Districts: An Ecologic Study. *Journal of School Health*, *84*(8), 533–542. <http://doi.org/10.1111/josh.12176>
- Jaycox, L. H., Stein, B. D., Paddock, S., Miles, J. N. V., Chandra, A., Meredith, L. S., ... Burnam, M. A. (2009). Impact of Teen Depression on Academic, Social, and Physical Functioning. *Pediatrics*, *124*(4), e596–e605. <https://doi.org/10.1542/peds.2008-3348>
- Johnston, L. D., Delva, J., & O'Malley, P. M. (2007). Sports Participation and Physical Education in American Secondary Schools: Current Levels and Racial/Ethnic and Socioeconomic Disparities. *American Journal of Preventive Medicine*, *33*(Supplement), S195–S208. <http://doi.org/10.1016/j.amepre.2007.07.015>
- Kalantari, H.-A., & Esmaeilzadeh, S. (2016). Association between academic achievement and physical status including physical activity, aerobic and muscular fitness tests in adolescent boys. *Environmental Health and Preventive Medicine*, *21*(1), 27–33. <http://doi.org/10.1007/s12199-015-0495-x>

- Kamijo, K., Khan, N. A., Pontifex, M. B., Scudder, M. R., Drollette, E. S., Raine, L. B., ... Hillman, C. H. (2012). The Relation of Adiposity to Cognitive Control and Scholastic Achievement in Preadolescent Children. *Obesity*, *20*(12), 2406–2411. <http://doi.org/10.1038/oby.2012.112>
- Kamijo, K., Pontifex, M. B., O’Leary, K. C., Scudder, M. R., Wu, C.-T., Castelli, D. M., & Hillman, C. H. (2011a). The effects of an afterschool physical activity program on working memory in preadolescent children. *Developmental Science*, *14*(5), 1046–1058. <http://doi.org/10.1111/j.1467-7687.2011.01054.x>
- Kantomaa, M. T., Tammelin, T. H., Demakakos, P., Ebeling, H. E., & Taanila, A. M. (2010). Physical activity, emotional and behavioural problems, maternal education and self-reported educational performance of adolescents. *Health Education Research*, *25*(2), 368–379. <http://doi.org/10.1093/her/cyp048>
- Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, *3*, 1–27. <https://doi.org/10.1146/annurev.clinpsy.3.022806.091432>
- Kirkendall, D. R. (1986). Effects of physical activity on intellectual development and academic performance. In Lee, M., Eckert, H.M., & Stull, G.A. (Eds.), *Effects of physical activity on children: A special tribute to Mabel Lee* (pp. 49–63). Champaign IL: Human Kinetics.
- Kitamura, T., Mishina, M., & Sugiyama, H. (2003). Enhancement of neurogenesis by running wheel exercises is suppressed in mice lacking NMDA receptor epsilon 1 subunit. *Neuroscience Research*, *47*(1), 55–63.
- Kobilo, T., Liu, Q.-R., Gandhi, K., Mughal, M., Shaham, Y., & van Praag, H. (2011). Running is the neurogenic and neurotrophic stimulus in environmental enrichment. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *18*(9), 605–609. <http://doi.org/10.1101/lm.2283011>
- Krafft, C. E., Schwarz, N. F., Chi, L., Weinberger, A. L., Schaeffer, D. J., Pierce, J. E., ... McDowell, J. E. (2014). An 8-month randomized controlled exercise trial alters brain activation during cognitive tasks in overweight children. *Obesity*, *22*(1), 232–242. <http://doi.org/10.1002/oby.20518>
- Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The Validity of Self-Reported Grade Point Averages, Class Ranks, and Test Scores: A Meta-Analysis and Review of the Literature. *Review of Educational Research*, *75*(1), 63–82. <https://doi.org/10.3102/00346543075001063>
- Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, *108*(3), 677–692. <http://doi.org/10.1016/j.jecp.2010.11.001>

- LeBlanc, M. D., & Weber-Russell, S. (1996). Text integration and mathematical connections: A computer model of arithmetic word problem solving. *Cognitive Science*, 20(3), 357–407. [http://doi.org/10.1016/S0364-0213\(99\)80010-X](http://doi.org/10.1016/S0364-0213(99)80010-X)
- Lee, K., Ng, E. L., & Ng, S. F. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology*, 101(2), 373–387. <http://doi.org/http://dx.doi.org.proxy1.cl.msu.edu/10.1037/a0013843>
- Lee, K., Ng, S. F., Pe, M. L., Ang, S. Y., Hasshim, M. N. A. M., & Bull, R. (2012). The cognitive underpinnings of emerging mathematical skills: Executive functioning, patterns, numeracy, and arithmetic. *British Journal of Educational Psychology*, 82(1), 82–99. <http://doi.org/10.1111/j.2044-8279.2010.02016.x>
- Lee, S., Burgeson, C., Fulton, J. E., & Spain, Christine G. (2007). Physical Education and Physical Activity: Results From the School Health Policies and Programs Study 2006. *Journal of School Health*, 77(8), 435–463.
- Lees, C., & Hopkins, J. (2013). Effect of Aerobic Exercise on Cognition, Academic Achievement, and Psychosocial Function in Children: A Systematic Review of Randomized Control Trials. *Preventing Chronic Disease*, 10. <http://doi.org/10.5888/pcd10.130010>
- Leung, L.-Y., Tong, K.-Y., Zhang, S.-M., Zeng, X.-H., Zhang, K.-P., & Zheng, X.-X. (2006). Neurochemical effects of exercise and neuromuscular electrical stimulation on brain after stroke: a microdialysis study using rat model. *Neuroscience Letters*, 397(1–2), 135–139. <http://doi.org/10.1016/j.neulet.2005.12.006>
- Lindner, K. J. (2002). The Physical Activity Participation-Academic Performance Relationship Revisited: Perceived and Actual Performance and the Effect of Banding (Academic Tracking). *Pediatric Exercise Science*, 14(2), 155–169.
- Lu, B., & Chow, A. (1999). Neurotrophins and hippocampal synaptic transmission and plasticity. *Journal of Neuroscience Research*, 58(1), 76–87.
- Markland, D., & Vansteenkiste, M. (2007). Self-Determination Theory and Motivational Interviewing in Exercise. In M. S. Hagger & N. L. D. Chatzisarantis, *Intrinsic Motivation and Self-Determination in Exercise and Sport* (1 edition). Champaign, IL: Human Kinetics.
- McGrew, K. S., & Woodcock, R. J. (2001). *Technical Manual: Woodcock-Johnson III*. Itasca, IL: Riverside.
- Miyake, A. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, 41(1), 49–100. <http://doi.org/10.1006/cogp.1999.0734>

- Molnar, B. E., Gortmaker, S. L., Bull, F. C., & Buka, S. L. (2004). Unsafe to Play? Neighborhood Disorder and Lack of Safety Predict Reduced Physical Activity Among Urban Children and Adolescents. *American Journal of Health Promotion, 18*(5), 378–386. <http://doi.org/10.4278/0890-1171-18.5.378>
- Molteni, R., Ying, Z., & Gómez-Pinilla, F. (2002). Differential effects of acute and chronic exercise on plasticity-related genes in the rat hippocampus revealed by microarray. *The European Journal of Neuroscience, 16*(6), 1107–1116.
- Morales, J., Pellicer-Chenoll, M., García-Massó, X., Gómez, A., Gomis, M., & González, L.-M. (2011). Relation between physical activity and academic performance in 3rd year secondary education students. *Perceptual and Motor Skills, 113*(2), 539–546. <http://doi.org/10.2466/06.11.13.PMS.113.5.539-546>
- Moss, M. C., Scholey, A. B., & Wesnes, K. (1998). Oxygen administration selectively enhances cognitive performance in healthy young adults: a placebo-controlled double-blind crossover study. *Psychopharmacology, 138*(1), 27–33. <https://doi.org/10.1007/s002130050641>
- Mulder, H., & Cragg, L. (Eds.). (2014). Executive Functions and Academic Achievement: Current Research and Future Directions: Editorial. *Infant and Child Development, 23*(1), 1–3. <http://doi.org/10.1002/icd.1836>
- Murray, N. G., Garza, J. C., Diamond, P. M., Hoelscher, D., Kelder, S., & Stigler, M. (2008). Classroom physical activity and Stanford 10 scores among 3rd and 4th grade students in Texas with adaptive skills problems. Presented at the American College of Sports Medicine Annual Conference.
- Naglieri, J. A., & Das, J. P. (1997). *Cognitive Assessment System*. Itasca, IL: Riverside Publishing Company.
- Naglieri, J. A., & Kaufman, J. C. (2001). Understanding intelligence, giftedness and creativity using the PASS theory. *Roepers Review, 23*(3), 151–156. <http://doi.org/10.1080/02783190109554087>
- Neeper, S. A., Gómez-Pinilla, F., Choi, J., & Cotman, C. W. (1996). Physical activity increases mRNA for brain-derived neurotrophic factor and nerve growth factor in rat brain. *Brain Research, 726*(1–2), 49–56. [http://doi.org/10.1016/0006-8993\(96\)00273-9](http://doi.org/10.1016/0006-8993(96)00273-9)
- Nelson, M. C., & Gordon-Larsen, P. (2006). Physical Activity and Sedentary Behavior Patterns Are Associated With Selected Adolescent Health Risk Behaviors. *Pediatrics, 117*(4), 1281–1290. <http://doi.org/10.1542/peds.2005-1692>
- No Child Left Behind Act of 2001, 20 U.S.C. § 6319 (2011).

- O'Callaghan, R. M., Ohle, R., & Kelly, A. M. (2007). The effects of forced exercise on hippocampal plasticity in the rat: A comparison of LTP, spatial- and non-spatial learning. *Behavioural Brain Research*, *176*(2), 362–366. <http://doi.org/10.1016/j.bbr.2006.10.018>
- Pan, L., Sherry, B., Park, S., & Blanck, H. M. (2013). The Association of Obesity and School Absenteeism Attributed to Illness or Injury Among Adolescents in the United States, 2009. *Journal of Adolescent Health*, *52*(1), 64–69. <http://doi.org/10.1016/j.jadohealth.2012.04.003>
- Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology*, *80*(1), 44–57. <http://doi.org/10.1006/jecp.2000.2626>
- Pate, R. R., Freedson, P. S., Sallis, J. F., Taylor, W. C., Sirard, J., Trost, S. G., & Dowda, M. (2002). Compliance with Physical Activity Guidelines: Prevalence in a Population of Children and Youth. *Annals of Epidemiology*, *12*(5), 303–308. [http://doi.org/10.1016/S1047-2797\(01\)00263-0](http://doi.org/10.1016/S1047-2797(01)00263-0)
- Pate, R. R., Mitchell, J. A., Byun, W., & Dowda, M. (2011). Sedentary behaviour in youth. *British Journal of Sports Medicine*, *45*(11), 906–913. <http://doi.org/10.1136/bjsports-2011-090192>
- Pate, R. R., & O'Neill, J. R. (2009). After-school interventions to increase physical activity among youth. *British Journal of Sports Medicine*, *43*(1), 14–18. <http://doi.org/10.1136/bjism.2008.055517>
- Plowman, S. A., & Meredith, M. A. (Eds.). (2013). *Fitnessgram/Activitygram Reference Guide* (4th ed.). Dallas Texas: The Cooper Institute.
- Pontifex, M. B., Kamijo, K., Scudder, M. R., Raine, L. B., Khan, N. A., Hemrick, B., ... Hillman, C. H. (2014). V. the Differential Association of Adiposity and Fitness with Cognitive Control in Preadolescent Children. *Monographs of the Society for Research in Child Development*, *79*(4), 72–92. <http://doi.org/10.1111/mono.12131>
- Pontifex, M. B., Raine, L. B., Johnson, C. R., Chaddock, L., Voss, M. W., Cohen, N. J., ... Hillman, C. H. (2011). Cardiorespiratory fitness and the flexible modulation of cognitive control in preadolescent children. *Journal of Cognitive Neuroscience*, *23*(6), 1332–1345. <http://doi.org/10.1162/jocn.2010.21528>
- Psychological Testing Corporation. (1990). *Metropolitan Achievement Tests, 6th ed.* San Diego, CA: Harcourt, Brace, Jovanovich.
- Reed, J. A., Einstein, G., Hahn, E., Hooker, S. P., Gross, V. P., & Kravitz, J. (2010). Examining the impact of integrating physical activity on fluid intelligence and academic performance in an elementary school setting: a preliminary investigation. *Journal of Physical Activity & Health*, *7*(3), 343–351.

- Reinert, K. R. S., Po'e, E. K., & Barkin, S. L. (2013). The Relationship between Executive Function and Obesity in Children and Adolescents: A Systematic Literature Review. *Journal of Obesity*, 2013. <http://doi.org/10.1155/2013/820956>
- Rennie, K. L., Livingstone, M. B. E., Wells, J. C., McGloin, A., Coward, W. A., Prentice, A. M., & Jebb, S. A. (2005). Association of physical activity with body-composition indexes in children aged 6–8 y at varied risk of obesity. *The American Journal of Clinical Nutrition*, 82(1), 13–20.
- Robbins, L. B., Pfeiffer, K. A., Vermeesch, A., Resnicow, K., You, Z., An, L., & Wesolek, S. M. (2013). “Girls on the Move” intervention protocol for increasing physical activity among low-active underserved urban girls: a group randomized trial. *BMC Public Health*, 13(1), 1–12. <http://doi.org/10.1186/1471-2458-13-474>
- Sallis, J. F., McKenzie, T. L., Kolody, B., Lewis, M., Marshall, S., & Rosengard, P. (1999). Effects of health-related physical education on academic achievement: project SPARK. *Research Quarterly for Exercise and Sport*, 70(2), 127–134.
- Sallis, J. F., Zakarian, J. M., Hovell, M. F., & Hofstetter, C. R. (1996). Ethnic, socioeconomic, and sex differences in physical activity among adolescents. *Journal of Clinical Epidemiology*, 49(2), 125–134.
- Sanchez, E., & Buddin, R. (2016). *How Accurate are Self-Reported High School Courses, Course Grades, and Grade Point Average?* (Research Report Series No. 2). ACT.
- Schinder, A. F., & Poo, M. (2000). The neurotrophin hypothesis for synaptic plasticity. *Trends in Neurosciences*, 23(12), 639–645.
- Schwarz, A. J., Brasel, J. A., Hintz, R. L., Mohan, S., & Cooper, D. M. (1996). Acute effect of brief low- and high-intensity exercise on circulating insulin-like growth factor (IGF) I, II, and IGF-binding protein-3 and its proteolysis in young healthy men. *The Journal of Clinical Endocrinology and Metabolism*, 81(10), 3492–3497. <http://doi.org/10.1210/jcem.81.10.8855791>
- Scudder, M. R., Lambourne, K., Drollette, E. S., Herrmann, S. D., Washburn, R. A., Donnelly, J. E., & Hillman, C. H. (2014). Aerobic capacity and cognitive control in elementary school-age children. *Medicine and Science in Sports and Exercise*, 46(5), 1025–1035.
- Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., ... Gore, J. C. (2004). Development of left occipitotemporal systems for skilled reading in children after a phonologically- based intervention. *Biological Psychiatry*, 55(9), 926–933. <https://doi.org/10.1016/j.biopsych.2003.12.019>
- Shephard, R. J. (1996). Habitual Physical Activity and Academic Performance. *Nutrition Reviews*, 54(4), S32–S36. <http://doi.org/10.1111/j.1753-4887.1996.tb03896.x>

- Sigfusdottir, I. D., Kristjansson, A. L., & Allegrante, J. P. (2007). Health Behaviour and Academic Achievement in Icelandic School Children. *Health Education Research*, 22(1), 70–80.
- Singh, A., Uijtdewilligen, L., Twisk, J. W. R., van Mechelen, W., & Chinapaw, M. J. M. (2012). Physical activity and performance at school: a systematic review of the literature including a methodological quality assessment. *Archives of Pediatrics & Adolescent Medicine*, 166(1), 49–55. <http://doi.org/10.1001/archpediatrics.2011.716>
- Sirard, J. R., & Pate, R. R. (2001). Physical Activity Assessment in Children and Adolescents. *Sports Medicine*, 31(6).
- Stevens, T. A., To, Y., Stevenson, S. J., & Lochbaum, M. R. (2008). The Importance of Physical Activity and Physical Education in the Prediction of Academic Achievement. *Journal of Sport Behavior*, 31(4), 368–388.
- Strauss, R. S. (1999). Comparison of measured and self-reported weight and height in a cross-sectional sample of young adolescents. *International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*, 23(8), 904–908.
- Strong, W. B., Malina, R. M., Blimkie, C. J. R., Daniels, S. R., Dishman, R. K., Gutin, B., ... Trudeau, F. (2005). Evidence Based Physical Activity for School-age Youth. *The Journal of Pediatrics*, 146(6), 732–737. <https://doi.org/10.1016/j.jpeds.2005.01.055>
- Swanson, H. L., & Alloway, T. P. (2012). Working memory, learning, and academic achievement. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook, Vol 1: Theories, constructs, and critical issues* (pp. 327–366). Washington, DC, US: American Psychological Association.
- Tang, S. W., Chu, E., Hui, T., Helmeste, D., & Law, C. (2008). Influence of exercise on serum brain-derived neurotrophic factor concentrations in healthy human subjects. *Neuroscience Letters*, 431(1), 62–65. <http://doi.org/10.1016/j.neulet.2007.11.019>
- Taras, H., & Potts-Datema, W. (2005). Obesity and Student Performance at School. *The Journal of School Health*, 75(8), 291–5.
- Tomporowski, P. D., Davis, C. L., Miller, P. H., & Naglieri, J. A. (2008). Exercise and Children's Intelligence, Cognition, and Academic Achievement. *Educational Psychology Review*, 20(2), 111–131. <http://doi.org/10.1007/s10648-007-9057-0>
- Tomporowski, P. D., Lambourne, K., & Okumura, M. S. (2011). Physical activity interventions and children's mental function: An introduction and overview. *Preventive Medicine: An International Journal Devoted to Practice and Theory*, 52(Suppl), S3–S9. <http://doi.org/10.1016/j.ypmed.2011.01.028>



- Trejo, J. L., Carro, E., & Torres-Aleman, I. (2001). Circulating insulin-like growth factor I mediates exercise-induced increases in the number of new neurons in the adult hippocampus. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *21*(5), 1628–1634.
- Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise*, *40*(1), 181–188. <http://doi.org/10.1249/mss.0b013e31815a51b3>
- Trost, S. G. (2007). State of the Art Reviews: Measurement of Physical Activity in Children and Adolescents. *American Journal of Lifestyle Medicine*, *1*(4), 299–314. <http://doi.org/10.1177/1559827607301686>
- Trost, S. G., McIver, K. L., & Pate, R. R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine and Science in Sports and Exercise*, *37*(11 Suppl), S531-543.
- Trudeau, F., & Shephard, R. J. (2010). Relationships of Physical Activity to Brain Health and the Academic Performance of Schoolchildren. *American Journal of Lifestyle Medicine*, *4*(2), 138–150. <http://doi.org/10.1177/1559827609351133>
- Tuckman, B. W., & Hinkle, J. S. (1986). An experimental study of the physical and psychological effects of aerobic exercise on schoolchildren. *Health Psychology*, *5*(3), 197–207. <http://doi.org/10.1037/0278-6133.5.3.197>
- Tulsky, D. S., Carlozzi, N. E., Chevalier, N., Espy, K. A., Beaumont, J. L., & Mungas, D. (2013). V. NIH TOOLBOX COGNITION BATTERY (CB): Measuring Working Memory. *Monographs of the Society for Research in Child Development*, *78*(4), 70–87. <http://doi.org/10.1111/mono.12035>
- U.S. Department of Health and Human Services. (2008, October). 2008 Physical Activity Guidelines for Americans. Retrieved from <http://health.gov/paguidelines/pdf/paguide.pdf>
- U.S. Department of Health and Human Services. (2010). *The association between school-based physical activity, including physical education, and academic performance*. Atlanta Georgia.
- U.S. Department of Health and Human Services, Centers for Disease Control, National Center for Chronic Disease Prevention and Health Promotion, & The President’s Council on Physical Fitness and Sports. (1996). *Physical Activity and Health: A Report of the Surgeon General*.
- van der Borght, K., Ferrari, F., Klauke, K., Roman, V., Havekes, R., Sgoifo, A., ... Meerlo, P. (2006). Hippocampal cell proliferation across the day: increase by running wheel activity, but no effect of sleep and wakefulness. *Behavioural Brain Research*, *167*(1), 36–41. <http://doi.org/10.1016/j.bbr.2005.08.012>

- van der Sluis, S., de Jong, P. F., & Leij, A. van der. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *Journal of Experimental Child Psychology*, 87(3), 239–266. <http://doi.org/10.1016/j.jecp.2003.12.002>
- van Praag, H. (2008). Neurogenesis and exercise: past and future directions. *Neuromolecular Medicine*, 10(2), 128–140. <http://doi.org/10.1007/s12017-008-8028-z>
- van Praag, H. (2009). Exercise and the brain: something to chew on. *Trends in Neurosciences*, 32(5), 283–290. <http://doi.org/10.1016/j.tins.2008.12.007>
- van Praag, H., Kempermann, G., & Gage, F. H. (1999). Running increases cell proliferation and neurogenesis in the adult mouse dentate gyrus. *Nature Neuroscience*, 2(3), 266–270. <http://doi.org/10.1038/6368>
- van Praag, H., Shubert, T., Zhao, C., & Gage, F. H. (2005). Exercise Enhances Learning and Hippocampal Neurogenesis in Aged Mice. *The Journal of Neuroscience*, 25(38), 8680–8685. <http://doi.org/10.1523/JNEUROSCI.1731-05.2005>
- Vaynman, S., Ying, Z., & Gomez-Pinilla, F. (2004). Hippocampal BDNF mediates the efficacy of exercise on synaptic plasticity and cognition. *The European Journal of Neuroscience*, 20(10), 2580–2590. <http://doi.org/10.1111/j.1460-9568.2004.03720.x>
- Voss, M. W., Nagamatsu, L. S., Liu-Ambrose, T., & Kramer, A. F. (2011). Exercise, brain, and cognition across the life span. *Journal of Applied Physiology (Bethesda, Md.: 1985)*, 111(5), 1505–1513. <http://doi.org/10.1152/jappphysiol.00210.2011>
- Wang, Z., Patterson, C. M., & Hills, A. P. (2002). A comparison of self-reported and measured height, weight and BMI in Australian adolescents. *Australian and New Zealand Journal of Public Health*, 26(5), 473–478. <http://doi.org/10.1111/j.1467-842X.2002.tb00350.x>
- Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., ... Gershon, R. C. (2013). Cognition assessment using the NIH Toolbox. *Neurology*, 80. <http://doi.org/10.1212/WNL.0b013e3182872ded>
- Weintraub, S., Bauer, P. J., Zelazo, P. D., Wallner-Allen, K., Dikmen, S. S., Heaton, R. K., ... Gershon, R.C.. (2013). NIH Toolbox Cognition Battery (CB): Introduction and Pediatric Data. *Monographs of the Society for Research in Child Development*, 78(4), 1–15. <http://doi.org/10.1111/mono.12031>
- Welk, G. J., Jackson, A. W., Morrow, J. R., Haskell, W. H., Meredith, M. D., & Cooper, K. H. (2010). The association of health-related fitness with indicators of academic performance in Texas schools. *Research Quarterly for Exercise and Sport*, 81(3 Suppl), S16-23. <http://doi.org/10.1080/02701367.2010.10599690>
- Wells, J. C. K. (2005). Measuring body composition. *Archives of Disease in Childhood*, 91(7), 612–617. <http://doi.org/10.1136/adc.2005.085522>

- Woodcock, R. W., McGrew, K. S., & Matdher, N. (2001). Woodcock-Johnson® III NU tests of achievement. *Rolling Meadows, IL: Riverside Publishing*.
- Yeniad, N., Malda, M., Mesman, J., van IJzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences, 23*, 1–9. <http://doi.org/10.1016/j.lindif.2012.10.004>
- Yu, C. C. W., Chan, S., Cheng, F., Sung, R. Y. T., & Hau, K. (2006). Are physical activity and academic performance compatible? Academic achievement, conduct, physical activity and self-esteem of Hong Kong Chinese primary school children. *Educational Studies, 32*(4), 331–341. <http://doi.org/10.1080/03055690600850016>
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1*(1), 297–301. <http://doi.org/10.1038/nprot.2006.46>
- Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). NIH Toolbox Cognition Battery: Measuring Executive Function and Attention. *Monographs of the Society for Research in Child Development, 78*(4), 16–33. <http://doi.org/10.1111/mono.12032>
- Zoladz, J. A., Pilc, A., Majerczak, J., Grandys, M., Zapart-Bukowska, J., & Duda, K. (2008). Endurance training increases plasma brain-derived neurotrophic factor concentration in young healthy men. *Journal of Physiology and Pharmacology: An Official Journal of the Polish Physiological Society, 59 Suppl 7*, 119–132.